

SEARCH REQUEST FORM

Requestor's Name: _____ Serial Number: _____
Date: _____ Phone: _____ Art Unit: _____

Search Topic:

Please write a detailed statement of search topic. Describe specifically as possible the subject matter to be searched. Define any terms that may have a special meaning. Give examples or relevant citations, authors, keywords, etc., if known. For sequences, please attach a copy of the sequence. You may include a copy of the broadest and/or most relevant claim(s).

STAFF USE ONLY

Date completed: 09-05-03
Searcher: Bodony 24494
Terminal time: 20
Elapsed time: _____
CPU time: _____
Total time: 23
Number of Searches: _____
Number of Databases: 1

Search Site

____ STIC
____ CM-1
____ Pre-S

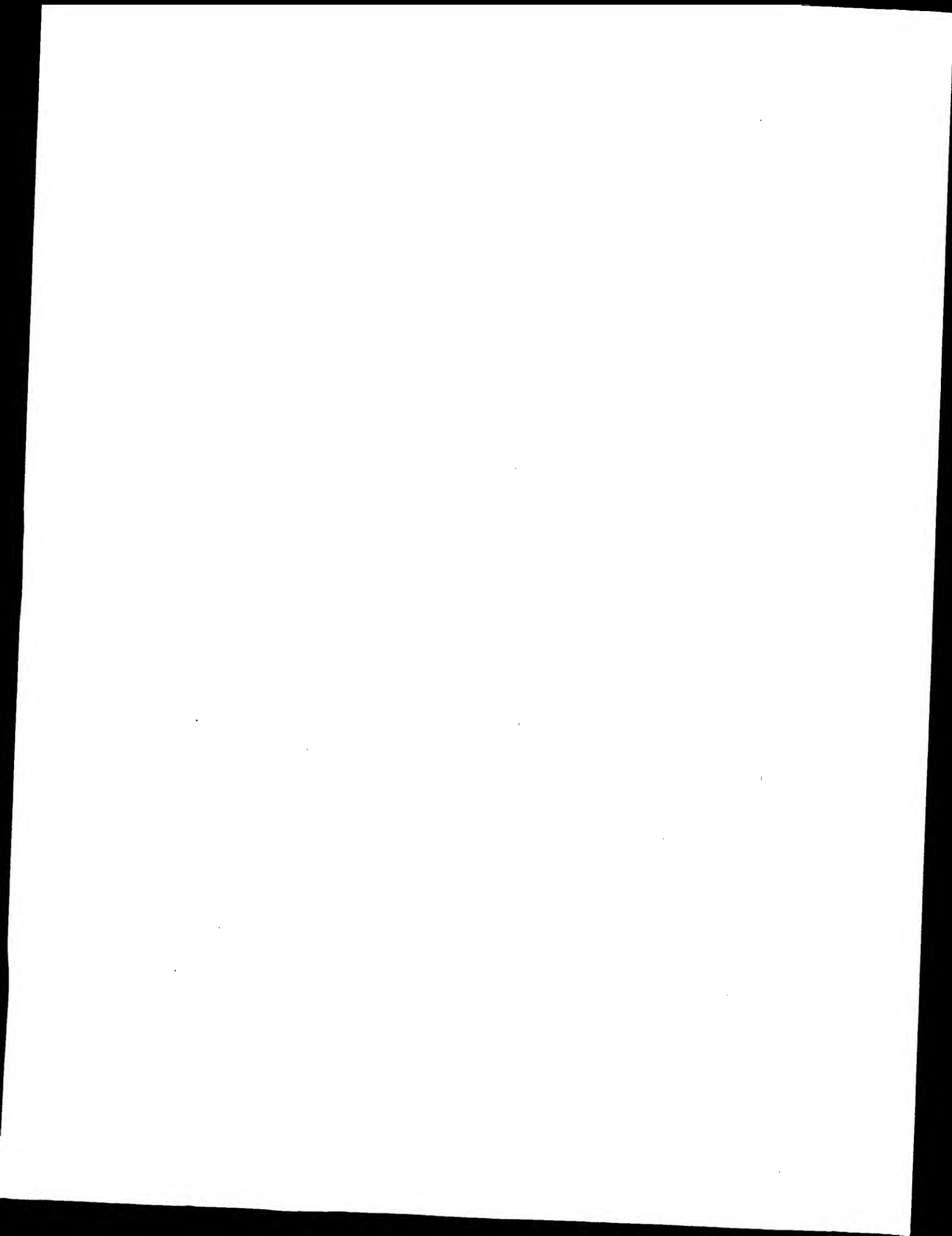
Type of Search

____ N.A. Sequence
____ A.A. Sequence
____ Structure
____ Bibliographic

Vendors

____ IG
____ STN
____ Dialog
____ APS
____ Geninfo
____ SDC
____ DARC/Questel
____ Other CGN

BEST AVAILABLE COPY



Fri Sep 5 09:48:11 2003

us-09-937-982-3.r1p11

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 4, 2003, 16:24:21 ; Search time 3324 Seconds

11169.064 Million cell updates/sec

Title: US-09-937-982-3
 Page: 1133

Sequence: 1 atgggcagcagccatcatca.....aagaaaggygaatctcttcac 1133

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 32997241 seqs, 16383922548 residues

Total number of hits satisfying chosen parameters: 65994482

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

Post-processing:	Minimum Match	0%
------------------	---------------	----

Listing first 45 summaries

Database : Pending_Patents_NA_main:*

1	/cgn2_6/ptodata/1/pna/pcrus.COMB.seq.*
2	/cgn2_6/ptodata/1/pna/pcrus.COMB.seq.old*
3	/cgn2_6/ptodata/1/pna/US06.COMB.seq.*
4	/cgn2_6/ptodata/1/pna/US07.COMB.seq.*
5	/cgn2_6/ptodata/1/pna/US08.COMB.seq.*
6	/cgn2_6/ptodata/1/pna/US082.COMB.seq.*
7	/cgn2_6/ptodata/1/pna/US082.COMB.seq.*
8	/cgn2_6/ptodata/1/pna/US083.COMB.seq.*
9	/cgn2_6/ptodata/1/pna/US083.COMB.seq.*
10	/cgn2_6/ptodata/1/pna/US085.COMB.seq.*
11	/cgn2_6/ptodata/1/pna/US086.COMB.seq.*
12	/cgn2_6/ptodata/1/pna/US087.COMB.seq.*
13	/cgn2_6/ptodata/1/pna/US088.COMB.seq.*
14	/cgn2_6/ptodata/1/pna/US089.COMB.seq.*
15	/cgn2_6/ptodata/1/pna/US090.COMB.seq.*
16	/cgn2_6/ptodata/1/pna/US091.COMB.seq.*
17	/cgn2_6/ptodata/1/pna/US092A.COMB.seq.*
18	/cgn2_6/ptodata/1/pna/US092B.COMB.seq.*
19	/cgn2_6/ptodata/1/pna/US093A.COMB.seq.*
20	/cgn2_6/ptodata/1/pna/US093B.COMB.seq.*
21	/cgn2_6/ptodata/1/pna/US094.COMB.seq.*
22	/cgn2_6/ptodata/1/pna/US095A.COMB.seq.*
23	/cgn2_6/ptodata/1/pna/US095B.COMB.seq.*
24	/cgn2_6/ptodata/1/pna/US095C.COMB.seq.*
25	/cgn2_6/ptodata/1/pna/US095D.COMB.seq.*
26	/cgn2_6/ptodata/1/pna/US096A.COMB.seq.*
27	/cgn2_6/ptodata/1/pna/US096B.COMB.seq.*
28	/cgn2_6/ptodata/1/pna/US096C.COMB.seq.*
29	/cgn2_6/ptodata/1/pna/US096C.COMB.seq.*
30	/cgn2_6/ptodata/1/pna/US096E.COMB.seq.*
31	/cgn2_6/ptodata/1/pna/US097A.COMB.seq.*
32	/cgn2_6/ptodata/1/pna/US097B.COMB.seq.*
33	/cgn2_6/ptodata/1/pna/US097C.COMB.seq.*
34	/cgn2_6/ptodata/1/pna/US098A.COMB.seq.*
35	/cgn2_6/ptodata/1/pna/US098B.COMB.seq.*
36	/cgn2_6/ptodata/1/pna/US098C.COMB.seq.*
37	/cgn2_6/ptodata/1/pna/US098D.COMB.seq.*
38	/cgn2_6/ptodata/1/pna/US099A.COMB.seq.*
39	/cgn2_6/ptodata/1/pna/US099C.COMB.seq.*
40	/cgn2_6/ptodata/1/pna/US099C.COMB.seq.*
41	/cgn2_6/ptodata/1/pna/US099D.COMB.seq.*
42	/cgn2_6/ptodata/1/pna/US099E.COMB.seq.*
43	/cgn2_6/ptodata/1/pna/US099F.COMB.seq.*

44:	/cgn2_6/p/odataa/1/pna/US1010A.COMB.seq.*
45:	/cgn2_6/p/odataa/1/pna/US1010B.COMB.seq.*
46:	/cgn2_6/p/odataa/1/pna/US1010C.COMB.seq.*
47:	/cgn2_6/p/odataa/1/pna/US1010E.COMB.seq.*
48:	/cgn2_6/p/odataa/1/pna/US1010Z.COMB.seq.*
49:	/cgn2_6/p/odataa/1/pna/US1010Z.COMB.seq.*
50:	/cgn2_6/p/odataa/1/pna/US1010A.COMB.seq.*
51:	/cgn2_6/p/odataa/1/pna/US1010B.COMB.seq.*
52:	/cgn2_6/p/odataa/1/pna/US1010C.COMB.seq.*
53:	/cgn2_6/p/odataa/1/pna/US1010E.COMB.seq.*
54:	/cgn2_6/p/odataa/1/pna/US6001.COMB.seq.*
55:	/cgn2_6/p/odataa/1/pna/US6001.COMB.seq.*
56:	/cgn2_6/p/odataa/1/pna/US6002.COMB.seq.*
57:	/cgn2_6/p/odataa/1/pna/US6003.COMB.seq.*
58:	/cgn2_6/p/odataa/1/pna/US6004.COMB.seq.*
59:	/cgn2_6/p/odataa/1/pna/US6005.COMB.seq.*
60:	/cgn2_6/p/odataa/1/pna/US6006.COMB.seq.*
61:	/cgn2_6/p/odataa/1/pna/US6007.COMB.seq.*
62:	/cgn2_6/p/odataa/1/pna/US6008.COMB.seq.*
63:	/cgn2_6/p/odataa/1/pna/US6009.COMB.seq.*
64:	/cgn2_6/p/odataa/1/pna/US6010.COMB.seq.*
65:	/cgn2_6/p/odataa/1/pna/US6011.COMB.seq.*
66:	/cgn2_6/p/odataa/1/pna/US6012.COMB.seq.*
67:	/cgn2_6/p/odataa/1/pna/US6013.COMB.seq.*
68:	/cgn2_6/p/odataa/1/pna/US6014.COMB.seq.*
69:	/cgn2_6/p/odataa/1/pna/US6015.COMB.seq.*
70:	/cgn2_6/p/odataa/1/pna/US6016.COMB.seq.*
71:	/cgn2_6/p/odataa/1/pna/US6017.COMB.seq.*
72:	/cgn2_6/p/odataa/1/pna/US6018.COMB.seq.*
73:	/cgn2_6/p/odataa/1/pna/US6019.COMB.seq.*
74:	/cgn2_6/p/odataa/1/pna/US6020.COMB.seq.*
75:	/cgn2_6/p/odataa/1/pna/US6021.COMB.seq.*
76:	/cgn2_6/p/odataa/1/pna/US6022.COMB.seq.*
77:	/cgn2_6/p/odataa/1/pna/US6023A.COMB.seq.*
78:	/cgn2_6/p/odataa/1/pna/US6023B.COMB.seq.*
79:	/cgn2_6/p/odataa/1/pna/US6024.COMB.seq.*
80:	/cgn2_6/p/odataa/1/pna/US6025.COMB.seq.*
81:	/cgn2_6/p/odataa/1/pna/US6026.COMB.seq.*
82:	/cgn2_6/p/odataa/1/pna/US6027.COMB.seq.*
83:	/cgn2_6/p/odataa/1/pna/US6028.COMB.seq.*
84:	/cgn2_6/p/odataa/1/pna/US6029.COMB.seq.*
85:	/cgn2_6/p/odataa/1/pna/US6030.COMB.seq.*
86:	/cgn2_6/p/odataa/1/pna/US6031.COMB.seq.*
87:	/cgn2_6/p/odataa/1/pna/US6032.COMB.seq.*
88:	/cgn2_6/p/odataa/1/pna/US6033.COMB.seq.*
89:	/cgn2_6/p/odataa/1/pna/US6034.COMB.seq.*
90:	/cgn2_6/p/odataa/1/pna/US6035.COMB.seq.*
91:	/cgn2_6/p/odataa/1/pna/US6036.COMB.seq.*
92:	/cgn2_6/p/odataa/1/pna/US6037.COMB.seq.*
93:	/cgn2_6/p/odataa/1/pna/US6038.COMB.seq.*
94:	/cgn2_6/p/odataa/1/pna/US6039.COMB.seq.*
95:	/cgn2_6/p/odataa/1/pna/US6040.COMB.seq.*
96:	/cgn2_6/p/odataa/1/pna/US6041.COMB.seq.*
97:	/cgn2_6/p/odataa/1/pna/US6042.COMB.seq.*
98:	/cgn2_6/p/odataa/1/pna/US6043.COMB.seq.*
99:	/cgn2_6/p/odataa/1/pna/US6044.COMB.seq.*
100:	/cgn2_6/p/odataa/1/pna/US6045.COMB.seq.*
101:	/cgn2_6/p/odataa/1/pna/US6046.COMB.seq.*
102:	/cgn2_6/p/odataa/1/pna/US6047.COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1133	100.0	1133	39	US-09-937-982-3	Sequence 3, Apple
2	1129.8	99.7	1133	33	US-09-773-280-3	Sequence 3, Apple
3	302.8	26.7	9082	37	US-09-897-516-1293	Sequence 1293, Apple
4	302.8	26.7	9082	37	US-09-897-516-1294	Sequence 1294, Apple

Malicka
09/1937982
Seq-ID 3 w/ Interf


```

: CURRENT APPLICATION NUMBER: US/09/897,516
: CURRENT FILING DATE: 2001-06-29
: PRIOR APPLICATION NUMBER: US 60/215, 161
: PRIOR FILING DATE: 2000-06-30
: NUMBER OF SEQ ID NOS: 8409
: SEQ ID NO 1293
: LENGTH: 9082
: TYPE: DNA
: ORGANISM: Xenorhabdus sp.
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (3968)..(5734)
: US-09-897-516-1293

```

```

Query Match      26.7%; Score 302.8; DB 37; Length 9082;
Best Local Similarity 58.1%; Pred. No. 3.4e-81;
Matches 576; Conservative 0; Mismatches 407; Indels 9; Gaps 2;

```

```

127 CCCCAAGTACTATCTAGCCACAGGAGGACCAATCCGCTGGTTCGGGGGAATCTAGCTC 186
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
3755 CCAAACTATTCTGCTGCGCAACGGGTGGCAATCGCTGGTGGTGAATCTGCAACT 3696
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
187 AAGAGTACTCTGCTGAGAGCACTGACCGTTGATAGCTTCTTGACCGCTCCGCTGC 246
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
3695 AAATCCAGCTACACTGGGGTCAAGTGGTGTGATTCATTGCTGAATGCTGCTCTCT 3636
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
247 ATCAACAGCTACGACCATTAAGGGTGAACAGATCTTCAGCAATGGCTCCCAAGATG 306
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
3635 ATCAAAAACATTTGTAATCTGTAAGAGTGACAGGTGTTTACATCGGTTCTCAGACAT 3576
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
307 ACGGTGAAGTGTGGCTTAACCTAGCCAGCGGTGATAGCTGCTCGCCCAAAAAGAG 366
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
3575 AAGCATGAGTGTGGCTGACTCTGCGCAAAAATCAATG-----CAGATTGATTA 3522
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
367 ACCGACCGCTGATCATACCATGGAATGAAAGAGACCGCTTCTCTC 426
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
3521 ACTGACGTTTGTATCATACCCAGGTACAGATACATGAAAGAACTGTTATTTCTC 3462
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
427 AACCTACGGTGAAGCAAAAACCTGCTGCTTGGAGCGCATCCGCTCCAGGCTCT 486
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
361 GATCTGACCACTGCTGATCAATACCTATCTAATGGTGGCAATGCTCCATCAACG 3402
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
487 TCATGAGTGTGATGAGCCCATGATCTATTAACCGCGTGAATGATGATCAACAA 546
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
3401 GCTCTGGGTGCTGATGGCCCATGATGACCTGATATGCTGTTGATGCGTCAATAAA 3342
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
547 GCGCTCTACTAACAAAGAGTGTGTGTGATGACATGATTCACGCCGCCAGAGAA 606
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
3341 GACTCTGCAAAACCGTGGTGTGTTTACTGCTATGATGATCTGTAATTCATGAGAC 3282
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
607 GCGACCAAGCTCAACACACACGCGAGTCAATGCTTGGCCCAACACAGGTAATATC 666
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
3281 ATGGGTAACTGAGCAACACTGAAGTTCAGGCTTCCAGGAGTGAACGCGCCCAAA 3222
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
667 GGCACACTCTATTATGCGAAAGTCGATTTTCACTCAATCCGTTGACCTCACACCTT 726
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
3221 GCGCTTCTATTATACGCTAAAGTAACTACTACTCT--GAGACCAAGGAAGAGAT 3165
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
727 GCAAGTGTGATATATTAACAAATGGAAGACTCCCGAGTCCATATCTTTAGCT 786
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
3164 AAGCTGTTTATGTCAGCAAACTGACTGAACCTGCAAAAGTATGTTATGTTATAC 3105
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
787 CACCCGAGTATGATGTTTATGTCATGAGCCCTTCAGGAGGAGGCAAGCAAGATC 846
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
3104 TACTCTAATGCTCCGATGCGCGGTTAAAGCATTTGTAGAGATGATTTTAAAGGAT 3045
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
847 ATCAACAGAGGATGAGCAATGGAACCTTTCCCTTGACTCAAAATGCTCTTGAAGAA 906
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
3044 ATCAGTGTGCTGTGTGTAACGCTAATATAGCTATCTGAGACCTCTGTGTACAA 2985
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
907 GAGGCAAAATCAGGCGTAGCGTCTCGAGCTTAGAGTGGGCACTGTTCCACACC 966
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2984 GCGGCTAAAGATGTGTGTGTGTTGTTCCAACCCGATTTCTTGGTTCACTACT 2925
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

QY 967 CAAGAGGCTGAAGTGGATGATAGAACTTGGTTTGGCTACAGAGTCTCAACCT 1026
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2924 CAAATGCTGAAGTAAAGACAGCAAAATCGGTTTGGTGGCAAGCGCTGAACCG 2865
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1027 CAAAAAGCCAGAGTCTCTTATGTTAGCCCTCACCAAAATAGATGAGAGAGGATC 1086
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2864 CAGAAAGCTGCTGTCTTCTTGTGCAATTTGGCTGACCTCAGACTGCTGACAGCAAAAT 2805
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1087 CAAAGATCTCTCCACCTATTATTCAGAA 1118
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2804 CAGGAATGTTTCTTAATTAATTCCTGAA 2773
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 4

```

US-09-897-516-1294/C
: Sequence 1294, Application US/09897516
: GENERAL INFORMATION:
: APPLICANT: Corbin, David R.
: APPLICANT: Goldman, Barry S.
: APPLICANT: Hinkle, Gregory J.
: APPLICANT: Huesing, Joseph E.
: APPLICANT: Krasomil-Osterfeld, Karina C.
: APPLICANT: Malvar, Thomas M.
: APPLICANT: Slater, Steven C.
: APPLICANT: Spiridonov, Sergei
: TITLE OF INVENTION: Xenorhabdus sp. Genome Sequences and Uses Thereof
: FILE REFERENCE: 38-21(51847)B
: CURRENT APPLICATION NUMBER: US/09/897,516
: CURRENT FILING DATE: 2001-06-29
: PRIOR APPLICATION NUMBER: US 60/215, 161
: PRIOR FILING DATE: 2000-06-30
: NUMBER OF SEQ ID NOS: 8409
: SEQ ID NO 1294
: LENGTH: 9082
: TYPE: DNA
: ORGANISM: Xenorhabdus sp.
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (6073)..(6924)
: US-09-897-516-1294

```

```

Query Match      26.7%; Score 302.8; DB 37; Length 9082;
Best Local Similarity 58.1%; Pred. No. 3.4e-81;
Matches 576; Conservative 0; Mismatches 407; Indels 9; Gaps 2;

```

QY 127 CCCCAAGTACTATCTAGCCACAGGAGGACCAATCCGCTGGTGGGGAATCTAGGCTC 186
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

DB 3755 CCAAACTATTCTGCTGCGCAACGGGTGGCAATCGCTGGTGGTGAATCTGCAACT 3696
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 187 AAGAGTACTCTGCTGAGAGCACTGACCGTTGATAGCTTCTTGACCGCTCCGCTGC 246
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

DB 3695 AAATCCAGCTACACTGGGGTCAAGTGGTGTGATTCATGATGATGCTGTTCTGCT 3636
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 247 ATCAAGACCTGACACCATCAAGGTGAAGATCTCAAGCATTTGGCTCCCAAGATG 306
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

DB 3635 ATCAAAAACATTTGTAATCTGTAAGAGTGACAGGTGTTGACATCGGTTCTCAGAGATG 3576
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 307 ACGGTGAAGTGTGGCTTAACCTAGCCAGCGTGTCAATGAGCTGCTCGCCCAAAAAGAG 366
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

DB 3575 AAGATGAGTGTGGCTGAGCTCTGCGCAAAAATCAATG-----CAGATTGATTA 3522
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 367 ACCGACCGCTGATCATACCATGGAAGTGAACCATGGAAGAGAGACCGCTTCTCTC 426
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

DB 3521 ACTGACGTTTGTATCATACCCAGGTACAGATACATGGAAGAACTGTTATTTCTC 3462
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 427 AACCTACGGTGAAGCAAAAACCTGCTGCTTGGAGCGCATCCGCTCCAGGCTCT 486
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

DB 3461 GATCTGACCACTGAGTCTCAATACCTATGTAATGGTGGCAATGCTCATCAACG 3402
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 487 TCATGAGTGTGATGAGCCCATGATCTATTAACCGCGTGAATGATGATCAACAA 546
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

DB 3401 GCTCTGGGTGCTGATGGCCCATGTAACCTGATATATGCTGTGTGAGGCTGATTA 3342
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

OY	547	GCCCTCTACTAACAAGAGATGGTGATGTGTAATGACCATGAATTCAGCGGCCCAAGAA	606
Db	3341	GACCTCCTCAAAACCCTGGTGTATTACTGCTGATGATGATATCTGTAAATTCATGACGCGAC	3282
OY	607	GCGAACAGGCCTCAACACACCACCGCACGTCAATGTCATTTGCTCGGCCAACACAGGTAAATC	666
Db	3281	ATCGGTAACTGACGACACACACTGAAAGTTCAAGCTTCAGGCGATGMAAGCAGGCGCCAA	3222
OY	667	GGCAGACTCTATTTATGCGAAAGTCGAGTATTTCACTCAATCCGTTGCACTCACACCTT	726
Db	3221	GGCTTCATCTAATACGGTAAAGTAACTACTACTCT--ACAGCACAGTGAAGAAGCAT	3165
OY	727	GCAAGTAGTTGGATATTAGCAAAATGAGAACTCCCCAGAGTCGATATTCTTACGCT	786
Db	3164	AAAGCTGTTTTTGGATGTCAGCAAACTAGTGAATGCGCAAAGATAGTATTGTGTATAAC	3105
OY	787	CACCCCGATGACTGATGTTTATGTCANTCAGACCCCTTACGCGAGACGCCAAAGCATC	846
Db	3104	TACTGTATGCTCCGATCTCCCGTTAAACATTTGTAGGATGATATTAAGGTATT	3045
OY	847	ATCCATGAGGCATAGGGCAATGGGACOCCTTCCCTTGAOTCCAATGCTCTGAAAA	906
Db	3044	ATCAGTGTGGTGTTGGTAACGGTBACATCTATAAGTCTATTCTGGACTCTCTTCACA	2985
OY	907	GCAGCCAAATCAGGCGTAGTCGTGCTCGTCGAAGCTCTAGAGTGGSCAGTGGTTCACACCC	966
Db	2984	GCGGCTAAAGATGGTGTGTGTGTGTTGCTTCCAAACGATATTCCTTCGGTTTCACTACT	2925
OY	967	CNAGSGCTGAAGTGATGATTAAGAACTTGSTTTGGCTTACAGAGCTCAACCT	1026
Db	2924	CAAAATGCTGAGTTAATGACGCAAAATACGTTTGTTCATCAGAACGCTGACACCG	2855
OY	1027	CAAAAAGCCAGAGTGTCTTATGCTTAGCCCTCACCAAAACTAGTATGATGAGAGCGATC	1086
Db	2864	CAGAAAGCTCGHGTCTTCTGCAATTTGGCTCTGACTGAGCTGACACAGCAAAAATT	2805
OY	1087	CANAAGTCTTCTCCACTATTAAATCCAAAA	1118
Db	2804	CAGGAATGTTTTCTAAATATTAAATTCCTGAA	2773
 RESULT 5 -US-09-897-516-1295/c Sequence 1295, Application US/09897516 GENERAL INFORMATION: APPLICANT: Corbin, David R. APPLICANT: Goldman, Barry S. APPLICANT: Hinkle, Gregory J. APPLICANT: Huesing, Joseph E. APPLICANT: Krasomil-Osterfeld, Karina C. APPLICANT: Malver, Thomas M. APPLICANT: Slater, Steven C. TITLE OF INVENTION: Xenorhabdus sp. Genome Sequences And Uses Thereof FILE REFERENCE: 38-21(51847)B CURRENT APPLICATION NUMBER: US/09/897,516 CURRENT FILING DATE: 2001-06-29 PRIOR APPLICATION NUMBER: US 60/215, 161 PRIOR FILING DATE: 2000-06-30 NUMBER OF SEQ ID NOS: 8409 SEQ ID NO 1295 LENGTH: 9082 TYPE: DNA ORGANISM: Xenorhabdus sp. FEATURE: NAME/KEY: CDS LOCATION: (6585)..(9081) US-09-897-516-1295			

Query Match	26.7%;	Score 302.8;	DB 37;	Length 9082;
Best Local Similarity	58.1%;	Pred. No. 3.4e-81;		
Matches 576; Conservative	0;	Mismatches 407;	Indels 9;	Gaps 2;

QY	127	CCCCAAGTACATCTCTAGCCACAGAGAGACCATCGCTGGTTCGGGGGAATCTACCGTC	186
Db	3755	CCAAACATCTTACTGTCTGGCAAGGGGGGACCAATCGCTGGTGGTGAATCTGCACT	3696
QY	187	AAGAGTAGCACTCGCTGGAGGAGTCACCGCTGGATAAAGCTTCTGACACCGCTCGCC	246
Db	3695	AAATCCAGCTCACTGCGGGGTCAAGTGGGTGGTATTCACTGGTGAATGCTGTCTCTGCT	3636
QY	247	ATCAACGACCTAGCCACCATCAAGGGTGAACAGATCTCAAGCATTTGGCTTCCACAGATG	306
Db	3635	ATCAAAAACATGTTACTATCTGAAGAGTGACAGAGTGGTATAGCATCGGTTCTCAGGACATG	3576
QY	307	ACGGGTAAAGTGTGGCTTAAACTTACGCCAACCGTGTCAATGAGCTCTCGCCAAAAAGG	366
Db	3575	AACGATCAGGCTGGCTGACTCTGGCCAAAAAAATCATG-----CAGATTGTGATAAA	3522
QY	367	ACCGAAGCCGATCATCATCCATCAGTAAGTACACATCGATGGAAGAGACCGCTTTCCTCTC	426
Db	3521	ACTGACGGTTTGTGATCATCCACCAAGTACAGATTCATGGAAGAACTGCTATTTCCTC	3462
QY	427	AACTCAGCGTGAAGAAAGCCAAAACCTGTCTGCTTTAGCGGCATAGCTTCCAGGCTCT	486
Db	3461	GATCTGACCACTCAGTGTGCATMAACATGCTGTAATGGTTGGCGAAGCGCTTCATCAACG	3402
QY	487	TCCATGATGTGCTGATGGGCCCATCAATCTCATTAACCCCTGATGAGCATCAACAAA	546
Db	3401	GCTTGGTGGCTGATGGGCCCATTAACCTCTATATGCTGTGTGTGAGGTCAGATATAA	3342
QY	547	GCCCTTACTTAACAAAGAGTGGTGTATGTGTGATGACATGAGATTCAAGCCGCCAGAGA	606
Db	3341	GACTGTGCAAAACCGGTGGTGTACTTCTGATGATATCTGTATATTCATGTGACGTGAC	3282
QY	607	GCGACAGCGTCACACACCGAGCGATCAATGCTATGGTGGCCACACAGGTAATATC	666
Db	3281	ATCGTAACTGAGCACAACCTGAAATTCAGGCTTCCAGGACGTAAGACAGGCCACAA	3222
QY	667	GGCACAGTCTATTATGGCAAGTCGAGATTTCACATCCGTTGCACTCACACCTT	726
Db	3221	GGCTTCATTCAATMAAGGTAAAGTAAACTACTACTCT--GCAGCACCAAGTAAAGCAGAT	3165
QY	727	GCAAGTAGTGTGATATTAGCAAAATCGAACAATCCCCACAGTCGATATCTTACGCT	786
Db	3164	AAAGGTGTTTTTGTGATGTCGCAAACTGACTACGACCCAAAAGTAGGTAATGTATTAC	3105
QY	787	CACCCGATGATACATGATGTTTATGTCATGACAGCCCTTCAGGACGACCCAAAGATC	846
Db	3104	TACTCTAATGCTTCCGATCTGCCGTTAAAGCATTTGTAGGAATGATTAATAAGGATTT	3045
QY	847	ATCCATGACGAGCGATGGGCATGGAACCTTCCCTTGACTCAAAATGCTCTTGA AAAA	906
Db	3044	ATCAGAGTGGTGGTGGTGAACGTAACCTAATTAAGTATATCTGAGACTCTCTGCACAA	2985
QY	907	GCAGCCAAATCAGGGGTAGTGGTGGCTCGTGAAGCTCTAGAGGGGACGTGGTTCACACC	966
Db	2984	GGCGCTAAAGATGAGTGGTGGTGGTGGTCTTCCAAACCGTATTTCTTGGGTTCACTACT	2925
QY	967	CAGAGGCTGAAGTGGATGATTAAGAACTTGGTTTGTGGCTACAGAGAGTCTCAACCT	1026
Db	2924	CAAAATGCTGAAGTAAATGACAGCAAAATACGGTTTTTGTGATCAGAAAGCGCTGACCCG	2865
QY	1027	CAAAAAGCCAGAGTGGTCTTATGTTAGGCCCTCACCAAAACATAGATAGAGAGGCGATC	1086
Db	2864	CAGAAAGCTCGGTCTTCTGCAATTTGGCTGACTCGACTGCTGACACAGCAAAATTT	2805
QY	1087	CAAAAGATCTTCCACCTATTAAATCCAGAA 1118	
Db	2804	CAGGAATGTTTTCTAAATATTAAATTATCCGAA 2773	

RESULT 6
US-09-897-516-1296
; Sequence 1296, Application US/09897516

OY		367	ACCGAAGCCGTAATCAATCACACCCTGGAAACGTGACACCATGGAGAAGAACCGCTTTCCTC	426
Db		5562	ACTGACGCTTTTGTCATCACCCACGGTAGAGATACATTGAAGAAACTGCTTATTTCTC	5621
OY		427	AACCTCAAGGTGAAAAGGCCAAAAAACCTGTGCCTGTGTAGCGCCATCGCTCCAGGCTCT	486
Db		5622	GATGTGACCACTAGTGTCATATAAACCTATGTAATGATGTGGCGAATGSCCTCCATCAACG	5681
OY		487	TCCATGAGTGTGTATGGCCCCCATGAAATCTTATPACGCCGTGAATGTAGGATCAACAA	546
Db		5582	GCTGTGGGTGTGATVGGCCCATTTGAACCTCTATPATCTCTGTGTAGSGTCAGATAAA	5741
OY		547	GCTCTACTAAACAAGAGATGATGTGTATGTATGATGACATGAGATTCAAGCCGCGACAGAA	606
Db		5742	GACTCTGCAAAACCGTGGTGTITTTACTTGTGATGATGATATCTGTAAATTCANMGACGTGAC	5801
OY		607	GGGACCAAGCTCAACACCACCGCAGTCAATGATCTTCTTGCCCAACACAGTAAATC	666
Db		5802	ATCGGTAACTGACACACACTGAAAGTTTCAGGCTTCCAGGACGTGAAGGCGCCCAA	5861
OY		667	GGCAGAGTCATTTATGGCAAGTGGAGATTTTCACTCAATCCGTTGCAACCTCACACCCTT	726
Db		5862	GGCTTCAATCTAATACGTTAAAGTAACTACTACTCT---CGAGCACAGTGAAGCAGAT	5918
OY		727	GCAAGTAGTTGTATTATTAGCAAAATCGAAGACTCCCAGAGTGCATTTCTTACGCT	786
Db		5919	AAAGCTGTTTTTGTATGTCAGCAAACTGACTGACMACTGCCAAAAGTAGATGTGTATPAC	5978
OY		787	CACCCCGATGATACMGATNGTTTATGTCOAATGACCCCTTCGAGGAGGCAAGGAATC	846
Db		5979	TACTCTAATGCTTCGATCTCCCGGTTTAAAGCATTTGTAGGAATGATATAAAGGTATT	6038
OY		847	ATCCATGCAGGATGGGCATGGGAACCCCTTCCCTTGTACTCAAAATGCTCTTGA AAAA	906
Db		6039	ATCATGCTGGTGTGTGAACGGTAACATCTATAAGTCTATTCTGSACTCTCTGCACAA	6098
OY		907	GCAGCCAATAGCGGTATGCTGGCTCGCTGSAAGCTCTAGAGGGGAGTGGTCCACACC	966
Db		6099	GCGGTAAAGATGGTGTGTGTGTGTGTTCCAAACGTATTTCTTTCGGTTTCACTACT	6158
OY		967	CAAGAGCGTGAAGTGATGATPAAGAAACTTGTTTTGTGGCTACAGAGAGTCAACCTT	1026
Db		6159	CAAAATGCTGAAGTATATGACAGCAAAATAGCGTTTTGTGCATCAGAAAGCCTGAACCG	6218
OY		1027	CAAAAAGCCAGAGTCTCTTATGTATTAGCCCTCACCAAAACTAGTATAGAGAGCGATC	1086
Db		6219	CAGAAAGCTGCGTCTTCTTCCAAATGGCTCTGACTGAGCTGCTGACACAGCAAAAATT	6278
OY		1087	CAAAAGATCTCTCCACCTATTATCTCAAGAA 1118	
Db		6279	CAGGAATGTTTTCTTAATATTATTTCTGAA 6310	
 RESULT 8 US-09-897-516-1298				
; Sequence 1298, Application US/09897516				
; GENERAL INFORMATION:				
; APPLICANT: Corbin, David R.				
; APPLICANT: Goldman, Barry S.				
; APPLICANT: Hinkle, Gregory J.				
; APPLICANT: Huesing, Joseph E.				
; APPLICANT: Krasomil-Osterfeld, Karina C.				
; APPLICANT: Malvar, Thomas M.				
; APPLICANT: Slater, Steven C.				
; APPLICANT: Spiridonov, Sergei				
; TITLE OF INVENTION: Xenorhabdus sp. Genome Sequences And Uses Thereof				
; FILE REFERENCE: 38-21(5184)B				
; CURRENT APPLICATION NUMBER: US/09/897,516				
; CURRENT FILING DATE: 2001-06-29				
; PRIOR APPLICATION NUMBER: US 60/715, 161				
; PRIOR FILING DATE: 2000-06-30				
; NUMBER OF SEQ ID NOS: 8409				
; SEQ ID NO 1298				

```

: LENGTH: 9082
: TYPE: DNA
: ORGANISM: Xenorhabdus sp.
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (5259)..(6302)
: US-09-897-516-1298

Query Match      26.7%; Score 302.8; DB 37; Length 9082;
Best Local Similarity 58.1%; Pred. No. 3,4e-81;
Matches 576; Conservative 0; Mismatches 407; Indels 9; Gaps 2;

QY      127  CCCCAAGTACTATCTCAGCCACAGGAGGACACATCGCTGTTGGGGGAATTAGCGTC 186
Db       5328  CCAAAACATTAATCTGTTGTGGCAACGGGAGGACATCGCTGGTGGTGAATCTGCACACT 5387
QY      187  AAGAGTAGCTCTCTGCTGGGAGCGTCACCGTTATTAAGGTTCTTGACAGCCGTCGCC 246
Db       5388  AAATTCACATACACTGGGAGGAGGAGTGGTGTATTTCATTGCTGGAATGCTGTTCCGCT 5447
QY      247  ATCAACGACCTAGCCACCATCAAGGGTGAACAGATCTCAAGCATTTGGCTCCCAAGATG 306
Db       5448  ATCAAAAACATTGCTTAATCTGAAGGTAGAGAGGTGTGTAGCATCGGTCTCCAGACATC 5507
QY      307  ACGGTAAAGTGTGGCTTAAACTGAGCCAGGCTGCATATAGACTCTCGCCCAAAAAGAG 366
Db       5508  AACGATCGGCTGGCTGACTCTGTGGCAAAAATCAATG-----CAGATTGTGATATA 5561
QY      367  ACCGAAAGCCGATCATCTACCCATGGAAGTACACATGAAAGAGACGGCTTCTTCCTC 426
Db       5562  ACTGACGGTTTTGTCTATCCACCGGTACAGATACCATGGAAGAACTGCTTATTTCTC 5621
QY      427  AACTCACAGGTGAAAAGCCAAAACCTGTCTGCTTGTAGGGCCGATCGCTCCAGGCTCT 486
Db       5622  GATCTGACCACTCACTGTCAATTAACCTATGCTTAATGTTGGCGAATGGCGCATCAACG 5681
QY      487  TCCATGATGTGCGATGGCCCATCATGATCTCTATAAGCCGCTGAATGTAGGCATCAACAA 546
Db       5682  GCTCTGGGTGCTGATGGCCCAATGAACTCTATATATGCTGTGTTAGGCTCAGATATA 5741
QY      547  GCGTCTACTAACAAAGAGTGGTGAATTTGTATGAACATGAGATTTACGCGCCAGAGAA 606
Db       5742  GACTCTGCAAAACCGTGGTGTTTTACTTCTATGATGATGTTCTGTAATTCATGACGTGAC 5801
QY      607  GCGACCAAGCTCAACACCAACCGACATCATGATTTGGTTGGCCACAGTAAATATC 666
Db       5802  ATCGGTAATCTAGCAACAATGTAAGTTAGGCTTCCACGAGAGTGAACGACGGCCACAA 5861
QY      667  GGCACAGTCTATTATGCAAGAGTGGATGATTTTCACTCAATCCGTTGCACTCACACCTT 726
Db       5862  GCGTTCATTCATAAGCGTAAAGTAAACTATCACTCT---GCAGACACCGTGAAGACGAT 5918
QY      727  GGAATGAGTTTGATATTAAGCAAAATCGAAGAACTCCCGAGAGTGATATCTTTACGCT 786
Db       5919  AAAAGCTTTTGAATGTCAGCAACTGACTGACACTGCCAAAAGTAGATTTGTGATTAAC 5978
QY      787  CACCCGAGTACTGATGTTTATGTCATGACGACCTTCAGGCGAGGACCAAGAAAGATC 846
Db       5979  TACTGTAAATGCTTCGATCTGCCGGTTAAAGCATTTGTAGAGAAATGATTAAGAGTAT 6038
QY      847  ATTCATGACAGGATGGGCAATGGGAACCTTCCCTTTGACTCAAAATGCTTTGAAAAA 906
Db       6039  ATCAGTGTGCTGTGTGGTAAACGTTAAACATCTTAAGTCTATTCTTGACCTCTCTGTCA 6098
QY      907  GCAGCCAAATTCAGGCGTAGTCTGCTCGTCGAAAGCTCTAGAGTGGGAGTGTTCACACC 966
Db       6099  GCGGCTTAAGATGCTGTGTTGTGTGCTTCCCAACCGATTCCTTTGGGTTTACACTCT 6158
QY      967  CAAAGAGCTGAAGTGAATGAAGAACTTGGTTTGTGGCTTACAGAGTCTCAACCTT 1028
Db       6159  CAAAATGCTGAAGTATATGACAGCAAAATCGGTTTGTGTCATCAGAAAGCGCTGAACCG 6218
QY      1027  CAAAAGCCAGAGTGGTTCTTATGTTAGGCCCTCACCAAAAACATGATGAGAGAGCGATC 1086

```

Db 6219 CAGAAAGCTGCTGTTCTTCTGCAATGCTGCTGACATGCTGACACAGCAAAAT 6278
 Oy 1087 CAAAAGATCTTCTCCACCTATTAATCCAGAA 1118
 Db 6279 CAGAAATGTTTCTTAATATTAATTCCTGAA 6310

RESULT 9 US-60-215-161-1293/c

Sequence 1293, Application US/60215161
 GENERAL INFORMATION:
 APPLICANT: Cordin, David R.
 APPLICANT: Goldman, Barry S.
 APPLICANT: Huesling, Joseph E.
 APPLICANT: Krasomil-Osterfeld, Karina C.
 APPLICANT: Malvar, Thomas M.
 APPLICANT: Slater, Steven C.
 APPLICANT: Spiridonov, Sergei
 APPLICANT: Hinkle, Gregory J.
 TITLE OF INVENTION: Xenorhabdus sp. Genome Sequences And Uses Thereof
 FILE REFERENCE: 38-21(51847)A
 CURRENT APPLICATION NUMBER: US/60/215,161
 CURRENT FILING DATE: 2000-06-30
 NUMBER OF SEQ ID NOS: 8409
 SEQ ID NO 1293
 LENGTH: 9082
 TYPE: DNA
 ORGANISM: Xenorhabdus sp.
 FEATURE:
 NAME/KEY: CDS
 LOCATION: (3868)..(5734)
 US-60-215-161-1293

Query Match 26.7%; Score 302.8; DB 75; Length 9082;
 Best Local Similarity 58.1%; Pred. No. 3,4e-81;
 Matches 576; Conservative 0; Mismatches 407; Indels 9; Gaps 2;

Oy 127 CCCCAGTACTTCTTCTAGCCACAGAGGACACATCGCTGCTGGGGATCTAGCGTC 186
 Db 3755 CCAAAATTAATCTTCTGCGACAGGGTGCGACAAATCGCTGGTGAATCGCAACT 3696
 Oy 187 AAGAGTAGTACTTCTGCTGAGCAGTCACCGTTGATTAAGCTCTGCGCGCCGCC 246
 Db 3695 AATTCAGCTACACTGCGGGTCAAGTGGTGTGATTCATTGCTGAATGCTGTCTGCT 3636
 Oy 247 ATACAGACTACCCACCATCAAGGAGGAGATCTCAAGATCTGCTCCCAAGAGATG 306
 Db 3635 ATCAAAAACATCTCTAATCTGAAGAGGAGCAGGTTGTAGCATCGGTTCTCAGGACATG 3576
 Oy 307 ACGGTAAGGTGGGTTAACTAAGCCACGCTGTAATGAGTCCCTCGCCCAAAAAGAG 366
 Db 3575 AACGATCAGGCTGCTGCTGACTGCGCAAAAATAATG-----CAGATTGTGATAAA 3522
 Oy 367 ACCGAAGCCGTGATCTACACCATGAGACTGACCATGAGAGAGACCGCTTCTTCTCTC 426
 Db 3521 ACTGACGGTTTGTCTATCAACCCAGGTAGCATGATGAGAAATACTGTTATTTCTC 3462
 Oy 427 AACCTACGCTGAAAAGCCAAAACCTGCTGCTGTGTAGCGCCATGCGTCCAGGCTCT 486
 Db 3461 GATCTGACCACTGATGATTAATTAACCTATGTAATGTTGGCCAAATGCGTCATCAAG 3402
 Oy 487 TCCATAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 546
 Db 3401 GCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3342
 Oy 547 GCTCTTACTAACAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 606
 Db 3341 GACTCTGCAACCGTGTGTTTACTTGTATGAATGATGATGATGATGATGATGATGATGAT 3282
 Oy 607 GGGACCAACCTCAACACCCAGATGATGATGATGATGATGATGATGATGATGATGATGAT 666
 Db 3281 ATCGGTAACGTAGCAACAGTGAAGTTTCAAGGCTTTCAGGAGTGAACGAGGCGCCACAA 3222

Oy 667 GGCACGCTATTAATGCAAAAGTGCAGATTTTCACTCAATCCGTTGCGATCACACCTT 726
 Db 3221 GCTTCATCTATTAACGATTAAGTAACTACTCT---GCAGCACGATGAAGACAGAT 3165
 Oy 727 GCAAGTGTGATTAATGCAAAATGCAAAAGTCCCAAGTGCATATCTTTAGCT 786
 Db 3164 AAAGCTGTTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3105
 Oy 787 CACCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 846
 Db 3104 TACTCTAATGCTTCCGATCTGCGGTTAAACATTTGATGATGATGATGATGATGATGATGAT 3045
 Oy 847 ATTCATGAGCAGATGAGGCAATGGAACCTTTCCCTTTGACTCAAAATGCTCTGAAAA 906
 Db 3044 ATCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2985
 Oy 907 GCACCAATATCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 966
 Db 2984 GCGGCTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2925
 Oy 967 CAAAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1026
 Db 2924 CAAATGCTGAAGTAAATGACAGCAAAATACGCTTTTGTGATGATGATGATGATGATGATGAT 2865
 Oy 1027 CAAAAGCGAGAGGCTTCTATGTTAGCCCTCACACCAAAATAGTATGATGATGATGATGATGAT 1086
 Db 2864 CAGAAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2805
 Oy 1087 CAAAAGATCTTCTCACCCTATTAATCCAGAA 1118
 Db 2804 CAGAAATGTTTCTTAATATTAATTCCTGAA 2773

RESULT 10

US-60-215-161-1294/c
 Sequence 1294, Application US/60215161

GENERAL INFORMATION:
 APPLICANT: Cordin, David R.
 APPLICANT: Goldman, Barry S.
 APPLICANT: Huesling, Joseph E.
 APPLICANT: Krasomil-Osterfeld, Karina C.
 APPLICANT: Malvar, Thomas M.
 APPLICANT: Slater, Steven C.
 APPLICANT: Spiridonov, Sergei
 APPLICANT: Hinkle, Gregory J.
 TITLE OF INVENTION: Xenorhabdus sp. Genome Sequences And Uses Thereof
 FILE REFERENCE: 38-21(51847)A
 CURRENT APPLICATION NUMBER: US/60/215,161
 CURRENT FILING DATE: 2000-06-30
 NUMBER OF SEQ ID NOS: 8409
 SEQ ID NO 1294
 LENGTH: 9082
 TYPE: DNA
 ORGANISM: Xenorhabdus sp.
 FEATURE:
 NAME/KEY: CDS
 LOCATION: (6073)..(6924)
 US-60-215-161-1294

Query Match 26.7%; Score 302.8; DB 75; Length 9082;
 Best Local Similarity 58.1%; Pred. No. 3,4e-81;
 Matches 576; Conservative 0; Mismatches 407; Indels 9; Gaps 2;

Oy 127 CCCCAGTACTTCTTCTAGCCACAGAGGACACATCGCTGCTGGGGATCTAGCGTC 186
 Db 3755 CCAAAATTAATCTTCTGCGACAGGGTGCGACAAATCGCTGGTGAATCGCAACT 3696
 Oy 187 AAGAGTAGTACTTCTGCTGAGCAGTCACCGTTGATTAAGCTCTGCGCGCCGCC 246
 Db 3695 AATTCAGCTACACTGCGGGTCAAGTGGTGTGATTCATGCTGAATGCTGTCTGCT 3636
 Oy 247 ATCAAGACCTAGCCACCATCAAGGTTGAACAGATCTCAAGCATTTGGTCCCAAGAGATG 306


```

Db      3635 ATCAAAAAATTGTAATCTGAAAGTGAAGCGAGTTGTTAGCATCGGTTCTCAGACATG 3576
      307  ACGGGTAAGTGTGGCTTAATTAAGTCAAGCGTCAATGAGTCCCTCGCCCAAAAGAG 366
      3575 AACGATCAGGTGTGGCTGACCTGTGGCAAAAAATCAATG-----CAGATTGTGATAA 3522
      367  ACCGAAGCCGTGATCATCAACCATGAACTGACACCATGGAAGAGACCGCTTCTCTC 426
      3521 ACTGACGGTTTTGTGATCAACCCAGCTACAGATACAGTAAGAAAGTCTATTCTCTC 3462
      427  AACCTCAGGGTGAAGAAAGCCAAACCTGCTCTGCTGAGGCGCATCGCTCAGGCTCT 486
      3461 GATCTGACACATCAGTGTATTAACCTATCGTAATGTTGGCGCATATGCTCATCAACG 3402
      487  TCCATGAGTGTGATGAGGCCCATGATCTCTATTAACCGCGTGAATGACGATCAACAA 546
      3401 GCTCTGGTGTGATGAGCCCATTTGAACCTATTAATGCTGTGTGTTGACGATATAA 3342
      547  GCCTCTACTACAAAGAGATGTTGATGATGAAGATGAGATTCACGCCCCAGAGAA 606
      3341 GACTCTGCAACCGTGTGTTTACTGCTATGAATGATTTGTAATTTATGAGCTGAC 3282
      607  GCGACCAAGCTCAACACCAACCGCATGATGCAATTTGCTGCGCCCAACAGTAAATC 666
      3281 ATCGGTAAGTGAACCAACCTGAAGTTCAAGGTTTCCAGGCAAGTGAACCGCCACAA 3222
      667  GGCACAGCTATTATGCAAGTGAAGTATTCACATTCGCTTGACCTCAGACCTT 726
      3221 GGGCTTATTCATCAACGTTAAGTAAACTACTACTCT---GCACGACCAAGTGAAGAGAT 3165
      727  GCAAGTGAAGTGAATTTGCAAAATCGAAAGACCTCCCAAGTCAATTTCTTTACGCT 786
      3164 AAAGCTGTTTTGATGTGACCAACTGACTGAACCTGCAAAAGTAGTATGTGTAAC 3105
      787  CACCCGATATCTGATGTTTATGTCATGCAATGCAAGCTTCAGGCGAGAGAGCAAGATC 846
      3104 TACTCTAATGCTTCCGATCTGCGGTTAAAGCATTTGTAGAGATGATTAAGAGTAT 3045
      847  ATCATCTCAGGCAATGGCAATGGGAATCCCTTCCCTTGACCTCAAAATGCTTTGAAAA 906
      3044 ATCATGCTGCTGTGTGTAACGTAACATCATTAATGCTATTTGCTGACTCTGTGACAA 2985
      907  GCACCAATATGAGCTGATGCTGCTGAGAGCTGATGAGTGGGCAAGTGTTCACACAC 966
      2984 GCGCTTAAGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 2925
      967  CAAAGAGCTGAAGTGAATAGAAATCTGTTTGTGCTACAGAGTCTCAACCT 1026
      2924 CAAATGCTGAAGTGAATAGAGCAATAGGTTTGTGCTACAGAGCTGTAACGCTGAACCG 2865
      1027 CAAAAGCCAGAGCTGCTTATATGTAGCCCTCACCAAAACCTAGTAGAGAGGCGATC 1086
      2864 CAGGAAGCTGCTGTTCTGCAATAGGCTGACTGACACTGCTGACACAGCAAAATTT 2805
      1087 CAAAAGATCTTCTCACCCTATTATTCAGAA 1118
      2804 CAGGAATGTTTCTAATATATTATTCCTGAA 2773

```

```

RESULT 11
US-60-215-161-1295/c
; Sequence 1295, Application US/60215161
; GENERAL INFORMATION:
; APPLICANT: Corbin, David R.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Huesling, Joseph E.
; APPLICANT: Krasomil-Osterfeld, Karina C.
; APPLICANT: Malvar, Thomas M.
; APPLICANT: Slater, Steven C.
; APPLICANT: Spiridonov, Sergei
; APPLICANT: Hinkle, Gregory J.
; TITLE OF INVENTION: Xenorhabdus sp. Genome Sequences And Uses Thereof

```

```

; FILE REFERENCE: 38-21(51847)A
; CURRENT APPLICATION NUMBER: US/60/215,161
; CURRENT FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 8409
; SEQ ID NO 1295
; LENGTH: 9082
; TYPE: DNA
; ORGANISM: Xenorhabdus sp.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (6985)..(9081)
; US-60-215-161-1295

Query Match      26.7%; Score 302.8; DB 75; Length 9082;
Best Local Similarity 58.1%; Pred. No. 3,4e-81;
Matches 576; Conservative 0; Mismatches 407; Indels 9; Gaps 2;

      127  CCCCAGTACTATCTTACGACAGAGGACCACTGCTGTTGGGGGATCAGGCTC 186
      3755  CCAACATTAAGTGTGTTGCAACGAGGTGCAAAATCGTGGTGTGATCTGCAACT 3696
      187  AAGATGACTCTCTGCTGAGCAGTCAACCTGATTAAGCTTCTGACGCGCTGCC 246
      3695  AATCCACTTACCTGCTGCGGTCMACTGCTGTTGATTCATTCGTAATGCTGTCGCT 3636
      247  ATCAAGCACTAAGCCACATCAAGGTGTAACAGATCTCAACATTTGCTCCCAAGATG 306
      3635  ATCAAAACATTTGTAATCTGAAAGTGAAGAGGATGTTTACATCGGTTCTCAGACATG 3576
      307  ACGGTGAAGTGTGGCTTAACCTAGCCAAAGGCTGTCATGAGCTCTCGCCAAAGAG 366
      3575  AACGATCAGGTGTGGCTGCTGCTGCGAATAAATCAATG-----CAGATTGTGATAA 3522
      367  ACCGAAGCCGTGATCATCAACCATGAACTGACACCATGGAAGAGACCGCTTCTCTC 426
      3521  ACTGACGGTTTTGTGATCAACCCAGCTACAGATACAGTAAGAAAGTCTATTCTCTC 3462
      427  AACCTCAGGGTGAAGAAAGCCAAACCTGCTCTGCTTGAAGCGCATCGCTCAGGCTCT 486
      3461 GATCTGACACATCAGTGTCTATTAACCTATGAAGTGTGGGCAAGTGTCTCATCAACG 3402
      487  TCCATGAGTGTGATGAGGCCCATGATCTCTATTAACCGCGTGAATGACGATCAACAA 546
      3401 GCTCTGGTGTGATGAGCCCATTTGAACCTCTTAATGCTGTGTGTTGAGCGTCAATATA 3342
      547  GCCTCTACTACAAAGAGATGTTGATGATGAAGATGAGATTCACGCCCCAGAGAA 606
      3341 GACTCTGCAACCGTGTGTTTACTGCTATGAATGATTTGTAATTCATGAGAGTGAAC 3282
      607  GCGACCAAGCTCAACACCAACCGCATGATGCAATTTGCTTGGCCCAACAGGTAATAATC 666
      3281 ATCGGTAAGCTGAGCAACATGAAGTTCAAGGCTTCCAGGCGAGTGAACGCCACAA 3222
      667  GGCACAGTCTATTAATGCAAAAGTCAGATTTTCACTCAATTCGTTGACCTCAACCTT 726
      3221 GGGCTTATTCATCAACGTTAAGTAAAGTGAAGCTTGTAGAGATGATTAAGAGTAT 3045
      727  GCAAGTGAAGTGAATATTAAGCAAAATCGAAGAACTCCCAAGAGTGAATTTCTTAAGCT 786
      3164 AAAGCTGTTTTGTGATGACCAAACTGACTGAACCTCACAAGTGAATGATTTGTAATAC 3105
      787  CACCCGATGATGATGATGTTTATGATCAATGACGCCCTCAGGCAAGAGCAAGAAAGATC 846
      3104 TACTCTAATGCTTCCGATCTGCGGTTAAAGCATTTGTGAGAGATATTTAAGGATAT 3045
      847  ATCATCTCAGGCAATGGCAATGGGAATCCCTTCCCTTGACCTCAAAATGCTTTGAAAA 906
      3044 ATCATGCTGCTGTGTGTAACGTAACATCATTAATGCTATTTGCTGACTCTGTACAA 2985
      907  GCACCAATATGAGCTGATGCTGCTGCTGCAAGCTCTAGAGTGGGCAAGTGTTCACACAC 966
      2984 GCGCTTAAGATGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 2925

```

OY	127	CCCCAAGGACATATCCTAGGCCACAGAGAGGCACATCGCTGGTTCGGGGGAATCTAGCGTC	186
Db	5328	CCAAACATTACTGTCTTGCGGCAACGGSTGGCACAAATCGTTGGTGGTGAATCTCAACT	5387
OY	187	AAGATGACTACTCTGCTGTGGAGCATGCACCGTGTGATAAGCTTCTTGACGCGCTCCCTCC	246
Db	5388	AAATTCAGCCTACACTGGGGGTCAACGTCCGTTGTATTATTGCTGAATAGCTGTCTCTCT	5444
OY	247	ATCAACGACCTAGGCCACCATCAGAAGGTGACAGATCTACAGCATGGCTCCCAAGAGATG	306
Db	5448	ATCAAAAAACATGCTTAATCTGAAAAGGTGAGCAGGTTGTACATCGGTTCTCAGACACTG	5507
OY	307	ACGGTAGAGGTGGCTTAAACTAGGCCAAGCGTGTCAATGAGCTCTCCGCCAAAAAGAG	366
Db	5508	AACGATCAGGCTGTGGCTGACTCTGTGGCGAAAAAACATCAATG-----CAGATTGTGATAAA	5561
OY	367	ACCGAAGCCCGATCATCACCATCGAGAACTGACACCATGGAAGAGACCGCTTTCTTCCTC	426
Db	5562	ACTACGCGTTTGTGATCATCCCAACGGTACAGATACCATGGAAGAAACGTGTTATTTCCTC	5621
OY	427	AACCTCAGCGTGAAGAAACCCAAAAACCTGTCTGCTGTAGAGCGGCATACCGTCAAGGCTCT	486
Db	5622	GATCTGACCACTCAGTGTGATTAACGTAATGTAATGTTGGCCGAATBGTCTCTCAACG	5681
OY	487	TCCATGAGTGCTGATGAGCCCAATGTAATCTATAVAGSCCGTAATGTAAGCATACAAGAA	546
Db	5682	GCTCTGGGTACTGTATGGCCCATTAAGCACTCTATAATGCTGTGTTGTAGAGCTCAGATAAA	5741
OY	547	GCCCTACTACCAAAGAGATGGTGTATGTATGACATGAGATTCAACGCCGCCAGAGAA	606

QY 127 CCCCAAGGACCTATCCCTAGCCACAGGAGGCCACCTCGCGGTTCGGGGGAATCTACGCCCTC 186
 Db 5328 CCAAACATTACTTCTCTGCGACAGCGGTGSCACAACTCGCTGGTGGTGATATCTGCCAATC 538

QY 187 AAGATGCTACTCTGCTGAGCAGTACCGTTGATTAAGCTTCTTGACACCGCTCCGCC 246
 Db 5388 AAATCCGCTACACTGCGGGGTCAAGTGGTGTATCATCTGTAATGCTGTTCTGCT 5447
 QY 247 ATCAACACCTAGCCACCATCAAGGTGAACAGATCCACCATTTGGCTCCAGAGATG 306
 Db 5448 ATCAAAAACATTTGCTTAATCTGAAGAGTGAAGAGTGTGTACCATGGTCTCAGAGACTG 5507
 QY 307 ACGGGTAAGGTGGCTTAACTAAGTCCAGGCGGTCTCAATGATGCTCTGCCCAAAAAGAG 366
 Db 5508 AAGCATGAGGTGCTGCTGACTGCTGCGCAAAAAATCAATG-----CAGATTGGATTA 5561
 QY 367 ACCGAGCCGTGATCATCAACCATGAGTACACCATGAGAGAGACCGCTTCTCTCTC 426
 Db 5562 ACTGACGGTTTGGTCAACCCAGGTACAGATACAGTGAAGAAAGTGGCTTATTTCTCTC 5621
 QY 427 AACCTCAAGTGAAGCAAAACCTGTGCTTGGAGGCGCATGCGCTCCAGGCT 486
 Db 5622 GATCTGACCACTGCTGCTATTAACCTATGCTAATGGTGGCGCAATGCTCCATCAAG 5681
 QY 487 TCCATGAGTGTGATGGCCCATGATATCTATTAACCGGTGATGAGTCAACAA 546
 Db 5682 GCTGTGGGTGTGATGGCCCATGATATCTATTAACCGGTGATGAGTCAACAA 5741
 QY 547 GCTCTACTAACAAGAGTGTGATGATGAGTCAAGTCAAGTCAAGCGCCAGAGAA 606
 Db 5742 GACTCTCAAAACCTGGTGTGCTGCTATGATGATGCTGATTAATCATGAGAGTAC 5801
 QY 607 GCGACAGGCTCAACACCGCAGTCAATGCTATTTGCTGCCCAACAGTAAATC 666
 Db 5802 ATCGGTAACTGACCACTGAGAGTCAAGCTTCCAGGAGTGAAGCGGCCCAAA 5861
 QY 667 GGCACAGTCTATTTGCAAAAGTCAAGTATTTCACTCAATCCGTTCCAGCTCACACCT 726
 Db 5862 GGCTTCAATCAACGTAAGTAAAGTAACTACTACTCT---GCAGCAGCTGAAGCAAT 5918
 QY 727 GCAAGTGAATTTAGTATTAAGCAAAATGCAAGTCAAGTCAAGTCAAGTCAAGT 786
 Db 5919 AAACCTGTTTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 5978
 QY 787 CACCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 846
 Db 5979 TACTCTAATGCTTCCGATCTGCGGTTAAAGCATTTGTAGAAATGATTAATTAAGTAT 6038
 QY 847 ATCCATCAGGATGAGGCAATGGAACCTTCCCTTGAATCAAAATGCTCTGAAAAA 906
 Db 6039 ATGATGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6098
 -QY 907 GCAGCCAAATCAGGCTGATGCTGCTGAGAGTCTAGAGTGGGCAAGTGGTCCACAC 966
 Db 6099 GCGGCTAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6158
 QY 967 CAAGAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1026
 Db 6159 CAATATGCTGAATTAATGACAGCAATACGTTTGTGATGATGATGATGATGATGATGAT 6218
 QY 1027 CAAAAACAGAGTCTTCTTATGTTAGTCCCTCAACAAATAGTATGATGATGATGATGAT 1086
 Db 6219 CAGAAAGTGTGCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6278
 QY 1087 CAAAAGATCTTCTCACCCTATTAATCAAGAA 1118
 Db 6279 CAGGAAATGTTTCTAATATTAATTAATCTCTGAA 6310

RESULT 14

US-60-215-161-1298
 ; Sequence 1298, Application US/60215161
 ; GENERAL INFORMATION:

; APPLICANT: Corbin, David R.
 ; APPLICANT: Goldman, Barry S.
 ; APPLICANT: Huesling, Joseph E.
 ; APPLICANT: Krasomil-Osterfeld, Karina C.

; APPLICANT: Malvar, Thomas M.
 ; APPLICANT: Slater, Steven C.
 ; APPLICANT: Spiridonov, Sergei
 ; APPLICANT: Hinkle, Gregory J.
 ; TITLE OF INVENTION: Xenorhabdus sp. Genome Sequences And Uses Thereof
 ; FILE REFERENCE: 38-21(51847)A
 ; CURRENT APPLICATION NUMBER: US/60/215,161
 ; CURRENT FILING DATE: 2000-06-30
 ; NUMBER OF SEQ ID NOS: 8409
 ; SEQ ID NO 1298
 ; LENGTH: 9082
 ; TYPE: DNA
 ; ORGANISM: Xenorhabdus sp.
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (5259)..(6302)
 ; US-60-215-161-1298

Query Match 26.7%; Score 302.8; DB 75; Length 9082;
 Best Local Similarity 58.1%; Pred. No. 3.4e-81;
 Matches 576; Conservative 0; Mismatches 407; Indels 9; Gaps 2;

QY 127 CCCAAGTACTATCTGAGCAGAGAGGACGACATCGCTGTTGGGGGATCTAGCGTC 186
 Db 5328 CCAACATTAATCTGTTGCGCAACGGGTGCAACATCGTGTGATGATGATGATGATGATGAT 5387
 QY 187 AAGATGCTACTCTGCTGAGCAGTACCGTTGATTAAGCTTCTTGACACCGCTCCGCC 246
 Db 5388 AAATCCGCTACACTGCGGGGTCAAGTGGTGTATCATCTGTAATGCTGTTCTGCT 5447
 QY 247 ATCAACACCTAGCCACCATCAAGGTGAACAGATCCACCATTTGGCTCCAGAGATG 306
 Db 5448 ATCAAAAACATTTGCTTAATCTGAAGAGTGAAGAGTGTGTACCATGGTCTCAGAGACTG 5507
 QY 307 ACGGGTAAGGTGGCTTAACTAAGTCCAGGCGGTCTCAATGATGCTCTGCCCAAAAAGAG 366
 Db 5508 AAGCATGAGGTGCTGCTGACTGCTGCGCAAAAAATCAATG-----CAGATTGGATTA 5561
 QY 367 ACCGAGCCGTGATCATCAACCATGAGTACACCATGAGAGAGACCGCTTCTCTCTC 426
 Db 5562 ACTGACGGTTTGGTCAACCCAGGTACAGATACAGTGAAGAAAGTGGCTTATTTCTCTC 5621
 QY 427 AACCTCAAGTGAAGCAAAACCTGTGCTTGGAGGCGCATGCGCTCCAGGCT 486
 Db 5622 GATCTGACCACTGCTGCTATTAACCTATGCTAATGGTGGCGCAATGCTCCATCAAG 5681
 QY 487 TCCATGAGTGTGATGGCCCATGATATCTATTAACCGGTGATGAGTCAACAA 546
 Db 5682 GGCTTCAATCAACGTAAGTAAAGTAACTACTACTCT---GCAGCAGCTGAAGCAAT 5918
 QY 547 GCTCTACTAACAAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 5978
 Db 5742 GACTCTCAAAACCTGGTGTGCTGCTATGATGATGATGATGATGATGATGATGATGATGAT 6038
 QY 607 GCGACAAATCAGGCTGATGCTGCTGAGAGTCTAGAGTGGGCAAGTGGTCCACAC 966
 Db 6099 GCGGCTAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6158
 QY 967 CAAGAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1026
 Db 6159 CAATATGCTGAATTAATGACAGCAATACGTTTGTGATGATGATGATGATGATGATGAT 6218
 QY 1027 CAAAAACAGAGTCTTCTTATGTTAGTCCCTCAACAAATAGTATGATGATGATGATGAT 1086
 Db 6219 CAGAAAGTGTGCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6278
 QY 1087 CAAAAGATCTTCTCACCCTATTAATCAAGAA 1118
 Db 6279 CAGGAAATGTTTCTAATATTAATTAATCTCTGAA 6310


```

Db      5508 AACGATCAGTCTGCTGACGCTGCGGCAAAAATCATG-----CAGATTGTGATAA 5561
QY      367 ACCGAAGCCGTGATCATCAACCATGGAAGTACACCATGGAAGAGACCCCTTTCTCC 426
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      5562 ACTGACGGTTTGTCTACACCCACGGTACAGATACAGTGAAGAACTCTATTCTTC 5621
QY      427 AACCTCAGGTGAAGGCAAAACCTGCTGCTGTGGGCGCCATGCTCAGGCTCT 486
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      5622 GATCTGACCACTCAGTCTATTAACCTATGCTATGATGTTGGCCATGCGTCCATCAAG 5681
QY      487 TCCATGAGTGTGATGAGCCCATGAATCTCTATTAACGCCGTGAATGATGATCAAAA 546
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      5682 GCTCTGGTGTGATGAGCCCATGACCTATTAATAGCTGTGTGTAGCTGACATGA 5741
QY      547 GCCTCTACTACAAAGAGTGGTGTGATGATGATGATGATGATGATGATGATGATGAT 606
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      5742 GACTCTGCAACCGTGTGTTTACTGCTATGAATGATCTGTAATTCATGAGACGTGAC 5801
QY      607 GCGACCAAGCTCAACCAACCGCATATGATGATGATGATGATGATGATGATGATGAT 666
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      5802 ATCGTAACTGACCACTGAAAGTTCAGGCTTCCAGGACGTGAAGCGAGGCCACAA 5861
QY      667 GGCACAGCTATTAATGCAAGTGTGATGATGATGATGATGATGATGATGATGATGAT 726
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      5862 GCTTCTATCATAGGTTAAAGTAACTACTACTCT---GCACACACAGTGAAGAGAT 5918
QY      727 GCAAGTATGTTGATTAATGAAGTAACTGCAAGTCCAGAGTGTATTTTACGCT 786
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      5919 AAACCTGTTTGTGATGACCAACTGACGTAAGTAACTGCAAGTCCAGAGTGTATTTTAC 5978
QY      787 CACCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 846
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      5979 TACCTATATGCTCGATCCGATCCGTTAAAGCATTTAGAGATGATTAAGAT 6038
QY      847 ATCCATGACGATGAGGCAATGGAAGTCCCTTCCCTTGAATGCTGGAAGAA 906
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      6039 ATCATGCTGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 6098
QY      907 GCAGCAATATGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 966
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      6099 GCGGCTTAAGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 6158
QY      967 CAAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1026
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      6159 CAATATGCTGAATTAATGACAGCAATATGCTGTTGTTGTTGTTGTTGTTGTT 6218
QY      1027 CAAAAGCCAGAGTCTCTTATGTTAGCTCCCAACCAAACTAGATGATGATGATGAT 1086
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      6219 CAGAAAGCTGCTGTTCTTCTGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCT 6278
QY      1087 CAAAAGATCTCTCTCCACTATTAATCAAGAA 1118
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      6279 CAGGAATGTTTCTTAATATTAATTAATTAATTAATTAATTAATTAATTAAT 6310

```

```

; NUMBER OF SEQ ID NOS: 8415
; SEQ ID NO 1297
; LENGTH: 9082
; TYPE: DNA
; ORGANISM: Xenorhabdus sp.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (6861)..(7958)
; OTHER INFORMATION:
US-09-897-516A-1297

Query Match      26.7%; Score 302.8; DB 5; Length 9082;
Best Local Similarity 58.1%; Pred. No. 2,7e-91;
Matches 576; Conservative 0; Mismatches 407; Indels 9; Gaps 2;

QY      127 CCCCAAGTCACTATCTTATGACCCACAGAGGACCACTATGCTGTTGGGGAAATCTAGCGTC 186
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      5328 CCAAACTATCTGTTCTGCAACGGGTGACAAATCGCTGGTGGTGAATCTGCAACT 5387
QY      187 AAGAGTAGTCTACTGCTGAGACAGTCAACCGTTGATTAAGCTTCTTGACCGCTCCCTGCC 246
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      5388 AATCCAGCTACACTGCGGGTCAAGTGGTGTGATGATGATGATGATGATGATGATGAT 5447
QY      247 ATCAAGACCTAGCCCACTCAAGAGGTGAACAGATCTCAAGCATGCTGCCCAAGAGATG 306
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      5448 ATCAAAACATTTGCTATCTGAAAGTGAAGAGTGTGATGATGATGATGATGATGATGAT 5507
QY      307 ACAGGTAGTGTGCTGTTAACTAGCCAAAGCTGTCAATGAGCTCTCGCCCAAAAGAG 366
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      5508 AACGATCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5561
QY      367 ACCGAAGCCGTGATCATCAACCATGGAAGTACACCATGGAAGAGACCCCTTTCTCC 426
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      5562 ACTGACGGTTTGTCTACACCCACGGTACAGATACAGTGAAGAACTCTATTCTTC 5621
QY      427 AACCTCAGGTGAAGGCAAAACCTGCTGCTGTGGGCGCCATGCTCAGGCTCT 486
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      5622 GATCTGACCACTCAGTCTATTAACCTATGCTATGATGTTGGCCATGCGTCCATCAAG 5681
QY      487 TCCATGAGTGTGATGAGCCCATGAATCTCTATTAACGCCGTGAATGATGATCAAA 546
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      5682 GCTCTGGTGTGATGAGCCCATGACCTATTAAGTCTGTTGTTGTTGTTGTTGTTGTT 5741
QY      547 GCCTCTACTACAAAGAGTGGTGTGATGATGATGATGATGATGATGATGATGATGAT 606
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      5742 GACTCTGCAACCGTGTGTTTACTGCTATGAATGATCTGTAATTCATGAGACGTGAC 5801
QY      607 GCGACCAAGCTCAACCAACCGCATATGATGATGATGATGATGATGATGATGATGAT 666
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      5802 ATCGTAACTGACCACTGAAAGTTCAGGCTTCCAGGACGTGAAGCGAGGCCACAA 5861
QY      667 GGCACAGCTATTAATGCAAGTGTGATGATGATGATGATGATGATGATGATGATGAT 726
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      5862 GCTTCTATCATAGGTTAAAGTAACTACTACTCT---GCACACACAGTGAAGAGAT 5918
QY      727 GCAAGTATGTTGATTAATGAAGTAACTGCAAGTCCAGAGTGTATTTTACGCT 786
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      5919 AAACCTGTTTGTGATGACCAACTGACGTAAGTAACTGCAAGTCCAGAGTGTATTTTAC 5978
QY      787 CACCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 846
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      5979 TACCTATATGCTCGATCCGATCCGTTAAAGCATTTAGAGATGATTAAGAT 6038
QY      847 ATCCATGACGATGAGGCAATGGAAGTCCCTTCCCTTGAATGCTGGAAGAA 906
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      6039 ATCATGCTGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 6098
QY      907 GCAGCAATATGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 966
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      6099 GCGGCTTAAGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 6158
QY      967 CAAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1026
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```


Db 6159 CAAATGCTGAAGTTAATGACAGCAAAATGCGTTTGGTTCATCAGAACGCCGTAACCCG 6218
 QY 1027 CAAAAGCCAGAGTCTTCTTATGTTAGCCCTCACCACAAATAGTATGAGAGGCGATC 1086
 Db 6219 CAGAAAGCTCGTCTTCTTCTGCAATTTGGCTGACCTCAGACGCTGACAGCAAAATTT 6278
 QY 1087 CAAAGATCTCTCCACCTATTAATCCAGAA 1118
 Db 6279 CAGGAATGTTTCTTAATTAATTAATTCCTGAA 6310

RESULT 7

US-09-897-516A-1298

Sequence 1298, Application us/09897516A

GENERAL INFORMATION:

APPLICANT: Corbin, David R.
 APPLICANT: Goldman, Barry S.
 APPLICANT: Hinkle, Gregory J.
 APPLICANT: Huesing, Joseph E.
 APPLICANT: Malvar, Thomas M.
 APPLICANT: Krasomil-Osterfeld, Karina C.
 APPLICANT: Slater, Steven C.
 APPLICANT: Splitdonov, Sergei
 TITLE OF INVENTION: Xenorhabdus sp. Genome Sequences And Uses Thereof
 FILE REFERENCE: 38-21(51847)B
 CURRENT APPLICATION NUMBER: US/09/897,516A
 CURRENT FILING DATE: 2001-06-29
 PRIOR APPLICATION NUMBER: US 60/215,161
 PRIOR FILING DATE: 2000-06-30
 NUMBER OF SEQ ID NOS: 8415
 SEQ ID NO 1298
 LENGTH: 9082
 TYPE: DNA
 ORGANISM: Xenorhabdus sp.
 FEATURE:
 NAME/KEY: CDS
 LOCATION: (5359)..(6302)
 OTHER INFORMATION:
 US-09-897-516A-1298

Query Match

Best Local Similarity 26.7%; Score 302.8; DB 5; Length 9082;

Matches 576; Conservative 0; Mismatches 407; Indels 9; Gaps 2;

QY 127 CCCCAGTACATCTCTAGCCACAGAGCACCACCTCGTGGGGGATCTAGAGCTC 186
 Db 5328 CCAACATCTACTGTTCTGCGAAGGGGCGACATCCCTGCTGTGTAATCTGCACT 5387
 QY 187 AAGAGTACTCTCTGTCGAGCAGTACCCGTTGATTAAGCTTTGCGAGCCCTGCC 246
 Db 5388 AATTCAGCTACACTCGGGTCAAGTCGGTGTGATTCATTTGCTGATGCTTCTGCT 5447
 QY 247 ATCAAGACCTAGCCACATCAAGGGAAGATCAAGTCAAGTGGTCCGCCAAGAGATG 306
 Db 5448 ATCAAAATATGCTATCTGAAGAGTGAACAGCTGTGATGCAATGCTTCCAGGATG 5507
 QY 307 ACGGGTAAAGTGTGCTTAATAGCAAGCGTGTCAATGAGCTCCCTGCCCAAAAGAG 386
 Db 5508 AAGATCAGGCTGCTGACTCTGGGAAAAAATCAATG-----CAGATTGTGATTA 5561
 QY 367 ACCGAAGCGGTGATCATCAACCATGAGTACACCATGGAAGAGACCGCTTCTTCCTC 426
 Db 5562 ACTGACGGTTTGTCTACACCAAGGTACAGATACCATGGAAGAACTCTTATTTCTC 5621
 QY 427 AACCTCAGCGTAAAAAGCCAAATCTGTCTCTTGTAGGCGCATGCGTCCAGGCTCT 486
 Db 5622 GATCTGACCATCTAGTGTATTAACCTATCGTAATGTTGGGCAATGCGCTCATTAAG 5681
 QY 487 TCCATGAGTGTGATGAGCCCAATGATCTTAAAGCCGATAGTACGATCAACAA 546
 Db 5682 GCTTGGGGTGTGATGAGCCATTAAGCTTAAATGCTGTGTGTACGTCAGATTA 5741
 QY 547 GCCTTACTAACAAGAGTGTGATTTGTGATGAAGATGAGATTCACGCCGCCAGAGA 606

Db 5742 GACTCTGCAAAACCGTGTCTTTTACTTGTATGAAGATTCGTATTAATCATGAGAGCTGAC 5801
 QY 607 GCGACCAAGCTCAACACCAAGCCAGTCAATGATTTGTTGGCCCAACAGGTAAATC 666
 Db 5802 ATCGTAACCTGAGCACAACCTAGTTCAGGCTTTCAGAGCAGTACAGCCACCA 5861
 QY 667 GGCACAGTATTAATGCGAAATCGAGTATTTCACTCAATCCGTTGACCTCAGACCTT 726
 Db 5862 GCTTCATTCATCAAGGTAAAGTAACTAACTACTCT--GCAGCACAGTGAACAGAT 5918
 QY 727 GCAGTGTATTAATTAATCAAAATCGAAGATCCCAAGTGTGATTAATCTTAACCT 786
 Db 5919 AAGCTGTTTGTGATGATGACAACTGACTGACTGACCAAAATGATTTGTATTAAC 5978
 QY 787 CACCCGATGATCTGATGTTTGTAGTCAATGACGCCCTTCAGGCGAGGACCAAGATC 846
 Db 5979 TACTCTAATGCTCCGATCTGCGGTAAAGCATTTGATGAAGATTAATTAAGTATT 6038
 QY 847 ATCCATGAGCATGCGCATGGAATGGGAAACCTTCCCTTGTACTCAAAATGCTTTGAAAA 906
 Db 6039 ATCACTGCTGCTGTTGTTGTAACGTAATCTATTAAGTCTATTTCTGACTCTCTGACAA 6098
 QY 907 GCAGCCAAATCAGGCGTATGCTGCTGCAAGCTTGAAGTGGGAGTGTCCACACAC 966
 Db 6099 GCGGCTAAAGATGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 6158
 QY 967 CAAGGCTGAAGTATGATTAAGAACTGTTTGTGCTGACAGAGTGTCAACCT 1026
 Db 6159 CAAATGCTGAAGTAAAGACAGCAAAATACGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 6218
 QY 1027 CAAAAGCCAGAGTCTTCTTATGTTAGCCCTCCACAAATCTAGTATGATGAGAGCGCATC 1086
 Db 6219 CAGAAAGCTGTGTTCTTCTCAATTTGGCTGTGACTGACTGCTGACACAGCAAAATTT 6278
 QY 1087 CAAAGATCTTCTCACCCTATTAATCCAGAA 1118
 Db 6279 CAGGAATGTTTCTTAATTAATTAATTCCTGAA 6310

RESULT 8

US-10-329-670-1

Sequence 1, Application US/10329670

GENERAL INFORMATION:

APPLICANT: Fletschmann et al.
 TITLE OF INVENTION: Nucleotide Sequence of the Haemophilus Influenzae Rd Genome, F
 FILE REFERENCE: P186P1
 CURRENT APPLICATION NUMBER: US/10/329,670
 PRIOR APPLICATION NUMBER: US 09/643,990
 PRIOR FILING DATE: 2000-08-23
 PRIOR APPLICATION NUMBER: US 08/487,429
 PRIOR FILING DATE: 1995-06-07
 PRIOR APPLICATION NUMBER: US 08/426,787
 PRIOR FILING DATE: 1995-04-21
 NUMBER OF SEQ ID NOS: 1
 SOFTWARE: Patent version 3.1
 SEQ ID NO 1
 LENGTH: 1830121
 TYPE: DNA
 ORGANISM: Haemophilus influenzae
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: (4747)..(4747)
 OTHER INFORMATION: n equals a, t, g or c
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: (9921)..(9921)
 OTHER INFORMATION: n equals a, t, g or c
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: (10150)..(10150)


```
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (29298)..(29298)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (36543)..(36543)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (36551)..(36551)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (36636)..(36636)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (40810)..(40810)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (44416)..(44416)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (44905)..(44905)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (44975)..(44975)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (45593)..(45593)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (45732)..(45732)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (47036)..(47036)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51334)..(51334)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51602)..(51602)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51786)..(51786)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51805)..(51805)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (55369)..(55369)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (65309)..(65309)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (65313)..(65313)
OTHER INFORMATION: n equals a, t, g or c
```

```
FEATURE:
NAME/KEY: misc_feature
LOCATION: (80024)..(80024)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (100091)..(100091)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (102696)..(102696)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (105121)..(105121)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (107248)..(107248)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (117136)..(117136)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (119750)..(119750)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (120038)..(120038)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (121344)..(121344)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (122167)..(122167)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (122336)..(122336)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (131340)..(131340)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (131360)..(131360)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (13910)..(13910)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (140398)..(140398)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (142750)..(142750)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145058)..(145058)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
```



```
SEQUENCE CHARACTERISTICS:
LENGTH: 966 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
HYPOTHEICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Streptococcus pneumoniae
FEATURE:
NAME/KEY: misc_feature
LOCATION: (B) LOCATION 1...966
SEQUENCE DESCRIPTION: SEQ ID NO: 791
US-10-617-320-791

Query Match
Best Local Similarity 51.9%; Score 84.4; DB 6; Length 966;
Matches 190; Conservative 0; Mismatches 176; Indels 0; Gaps 0;

QY 369 CGAAGCGGTATGATGACCATGAGACGACCATGAGAGAGCGCTTCTCTCA 428
DB 231 CGATGGAGGTGTCACACGACGACGACGACGACGACGACGACGACGACG 290
QY 429 CCTCACGCGTAAAAAGCCAAACCTGCTGCTGTAGCGCCGACGCTCCAGGCTTTC 488
DB 291 TACCATGGAGTCCCATATGCTATGCTTAAACAGAGGACGATGCTGCTCAATGA 350
QY 489 CATGATGCTGATGCGCCCATGATCTTATTAAGCGCCGCTGATGAGGATCAACAAAGC 548
DB 351 GCTCGGAGTATGATGCTTATTAATTAACCTTAAGCTTAAAGGCGCAAGTACAG 410
QY 549 CTCTACTAACAAGAGAGTGTATGATGATGATGATGATGATGATGATGATGATG 608
DB 411 GCGCTGACAAAGAGATTTGGTTCGTATGACATGATAAATCCAGCGTCCCAAGTAT 470
QY 609 GACCAAGCTCAACACGACGACGATGATGATGATGATGATGATGATGATGATG 668
DB 471 CACCAAAACATATGACATGATGATGATGATGATGATGATGATGATGATGATG 530
QY 669 CACAGCTATTTATGGAAGTGCATGATGATGATGATGATGATGATGATGATG 728
DB 531 TCTCATATGAAACAGAAATCTCTACTCAAAACAGCTGACCTGCTGCTTGA 590
QY 729 AAGTGA 734
DB 591 CCTTGA 596

RESULT 12
US-10-406-676-3
; Sequence 3, Application US/10406676
; GENERAL INFORMATION:
; APPLICANT: Structural Genomics, Inc.
; APPLICANT: Antonysamy, Stephen
; APPLICANT: Fell, Ingeborg
; APPLICANT: Buchanan, Sean
; APPLICANT: Post, Kai W.
; APPLICANT: Liu, Yi
; APPLICANT: Lober, David
; TITLE OF INVENTION: CRYSTALS AND STRUCTURES OF PAK4KD KINASE
; FILE REFERENCE: 524982002300
; CURRENT APPLICATION NUMBER: US/10/406,676
; PRIOR FILING DATE: 2003-04-02
; PRIOR APPLICATION NUMBER: 60/371,018
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/430,567
; PRIOR FILING DATE: 2002-12-02
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 82
```

```
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Sequence after ligation into vector
US-10-406-676-3

Query Match
Best Local Similarity 98.6%; Score 67.4; DB 6; Length 82;
Matches 68; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGGAGCAGCATCATCATCATCATCATCATCATCATCATCATCATCATCATCAT 60
DB 14 ATGGGAGCAGCAGCATCATCATCATCATCATCATCATCATCATCATCATCATCAT 73
QY 61 ATGGCTAGC 69
DB 74 ATGGCTAGC 82

RESULT 13
PCT-US03-06661A-9
; Sequence 9, Application PC/TUS0306661A
; GENERAL INFORMATION:
; APPLICANT: Chou, Szu-Yi
; TITLE OF INVENTION: Method of Producing Antigens
; FILE REFERENCE: SAMG/0002 PCT
; CURRENT APPLICATION NUMBER: PCT/US03/06661A
; CURRENT FILING DATE: 2003-03-03
; PRIOR APPLICATION NUMBER: 60/361,166
; PRIOR FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: 60/363,445
; PRIOR FILING DATE: 2002-03-08
; PRIOR APPLICATION NUMBER: 10/231,114
; PRIOR FILING DATE: 2002-08-28
; PRIOR APPLICATION NUMBER: 10/231,470
; PRIOR FILING DATE: 2002-08-28
; PRIOR APPLICATION NUMBER: 10/231,063
; PRIOR FILING DATE: 2002-08-28
; PRIOR APPLICATION NUMBER: 10/231,213
; PRIOR FILING DATE: 2002-08-28
; PRIOR APPLICATION NUMBER: 10/231,298
; PRIOR FILING DATE: 2002-08-28
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patent version 3.1
; SEQ ID NO 9
; LENGTH: 1067
; TYPE: DNA
; ORGANISM: Streptomyces mobaraensis ATCC 29032
PCT-US03-06661A-9

Query Match
Best Local Similarity 98.6%; Score 67.4; DB 1; Length 1067;
Matches 68; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGGAGCAGCATCATCATCATCATCATCATCATCATCATCATCATCATCATCAT 60
DB 1 ATGGGAGCAGCAGCATCATCATCATCATCATCATCATCATCATCATCATCATCAT 60
QY 61 ATGGCTAGC 69
DB 61 ATGGCTAGC 69

RESULT 14
PCT-US03-06661A-11
; Sequence 11, Application PC/TUS0306661A
; GENERAL INFORMATION:
; APPLICANT: Chou, Szu-Yi
; TITLE OF INVENTION: Method of Producing Antigens
; FILE REFERENCE: SAMG/0002 PCT
; CURRENT APPLICATION NUMBER: PCT/US03/06661A
; CURRENT FILING DATE: 2003-03-03
; PRIOR APPLICATION NUMBER: 60/361,166
```

Query Match	5.9%	Score 67.4	DB 1	Length 1074
Best Local	98.6%	Pred. No. 5.2e-12		
Matches 68	Conservative	0	Mismatches 1	Indels 0
				Gaps 0

RESULT 15
US-10-423-156-60
; Sequence 60, Application US/10423156
; GENERAL INFORMATION.

APPLICANT: Lin, Hsin-Yu
APPLICANT: Hwang, Ching-Long
TITLE OF INVENTION: ANTIGENIC FRAGMENT OF HUMAN
TITLE OF INVENTION: T-TYMPHOTROPIC VIRUS
FILE REFERENCE: 05204-020001
CURRENT APPLICATION NUMBER: US/10/423,156
CURRENT FILING DATE: 2003-04-25
PRIOR APPLICATION NUMBER: TW 911359802
PRIOR FILING DATE: 2002-12-12
NUMBER OF SEQ ID NOS: 60
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 60
LENGTH: 759
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Synthetically generated oligonucleotide
US-10-423-156-60

Query Match	5.84;	Score	65.2;	DB	6;	Length	759;
Best local	Similarity	95.7%;	Pred.	No.	2.4e11;		
Matches	67;	Conservative	0;	Mismatches	3;	Indels	0;
						Gaps	0

QY 1 ATGGGAGAGGCATCATCATCATCATCATATGACAGGCGCGTGTCCCGGGGCGACCAT 60
 |||||
 Db 364 ATGGGAGAGGCATCATCATCATCATCATCAGCAGGCGCGCTGTGTCCCGGGGAGGCAT 423
 |||||
 QY 61 ATGGCTAGCA 70
 |||||
 Db 424 ATGGCTTCCA 433

Search completed: September 4, 2003, 19:14:51
Job time : 217 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 4, 2003, 16:08:40 ; Search time 4197 seconds

(without alignments)
11043.747 Million cell updates/sec

Title: US-09-937-982-3

Perfect score: 1133

Sequence: 1 atgggcagcagccatcatca.....aagaagggatctcttcac 1133

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

```
Database :
1: gb_ba:*
2: gb_mt9:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pin:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_hgo_hum:*
40: em_hgo_mus:*
41: em_hgo_other:*
```

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1129.8	99.7	1133	6	ARI57944	ARI57944 Sequence
2	1129.8	99.7	1133	6	AX464427	AX464427 Sequence
3	1027.8	90.7	1035	1	WSANSAGEN	X83689 W.succinoge
4	1027.8	90.7	1035	1	WSDCUNSA	X89215 W.succinoge
5	368.8	32.6	3606	1	WSA12933	AJ002933 Wolinella
6	323.2	28.5	303450	1	AP005085	AP005085 Vibrio pa
7	300	26.5	10893	1	AE015308	AE015308 Shigella
8	300	26.5	292906	1	AE016988	AE016988 Shigella
9	296.8	26.2	10745	1	AE005526	AE005526 Escherich
10	296.8	26.2	266658	1	AP002563	AP002563 Escherich
11	296.8	26.2	303121	1	AE016766	AE016766 Escherich
12	296.4	26.2	10173	1	U32758	U32758 Haemophilus
13	296.4	26.2	110000	6	AR274513_07	Continuation (8 of
14	296.4	26.2	110000	6	AR274513_08	Continuation (9 of
15	295.2	26.1	1530	1	ECOLASNTI	M34277 E.coli L-as
16	295.2	26.1	1643	1	ECOLASNTI	M34277 E.coli L-as
17	295.2	26.1	10448	1	AE000378	M34234 E.coli L-as
18	295.2	26.1	141744	1	ECU28377	U28377 Escherichia
19	294	25.9	1848	6	I73512	I73512 Sequence 15
20	294	25.9	1848	12	U06943	U06943 Synthetic c
21	282.6	24.9	10769	1	AE013881	AE013881 Yersinia
22	282.6	24.9	210050	1	AJ414148	AJ414148 Yersinia
23	276	24.4	22204	1	AE008842	AE008842 Salmonell
24	276	24.4	230050	1	AL627277	AL627277 Salmonell
25	276	24.4	301574	1	AE016844	AE016844 Salmonell
26	273.2	24.1	305961	1	AE016937	AE016937 Bacteroid
27	270.4	23.9	10094	1	AE001498	AE001498 Helicobac
28	259.2	22.9	1196	6	BD092628	BD092628 Identific
29	259.2	22.9	10923	1	AE000585	AE000585 Helicobac
30	230.4	20.3	304500	1	AP005953	AP005953 Bradyrhiz
31	224.8	19.8	2450	1	ECSC	X14777 Erwinia chr
32	223.2	19.7	2837	6	A14577	A14577 asparaginase
33	223.2	19.7	2837	6	E01113	E01113 Genomic DNA
34	221	19.5	35000	1	AB000617	AB000617 Bacillus
35	221	19.5	221160	1	BSUB0002	Z99105 Bacillus su
36	215.2	19.0	1399	1	ECASN	X12746 Erwinia chr
37	210.8	18.6	1017	6	ARI77619	ARI77619 Sequence
38	209.2	18.5	1709	1	AF056495	AF056495 Pseudomon
39	207.6	18.3	314150	1	CJ11168X1	AI139074 Campyloba
40	205.2	18.1	308015	1	AE016783	AE016783 Pseudomon
41	203.2	17.9	110000	2	AC020884_1	Continuation (2 of
42	200.4	17.7	9937	1	AE004563	AE004563 Pseudomon
43	195.4	17.2	2353	1	PFL238710	AJ328710 Pseudomon
44	178.8	15.8	20516	1	AE008866	AE008866 Salmonell
45	128.8	11.4	1337	8	SPLASPA	Y11944 S.pombe gen

ALIGNMENTS

```
RESULT 1
LOCUS      ARI57944
DEFINITION Sequence 3 from patent US 6251388.
ACCESSION  ARI57944
VERSION     ARI57944.1 GI:16219888
KEYWORDS
SOURCE      Unknown.
ORGANISM    Unclassified.
REFERENCE   1 (bases 1 to 1133)
AUTHORS     Darden,D.L.
TITLE        Utilization of Wolinella succinogenes asparaginase to treat
             diseases associated with asparagine dependence
             Patent: US 6251388-A 3 26-JUN-2001;
JOURNAL
```


QY 823 CTTCAGGAGAGAGCCAAAGAAATCATCATGAGGATGGGCAATGGGAAACCTTTCCCT 882
 DB 725 CTTCAGGAGAGAGCCAAAGAAATCATCATGAGGATGGGCAATGGGAAACCTTTCCCT 784
 QY 883 TTGACTCAAAATGCTCTTGAAGAAAGAGCCAAATGAGGCTAGTCTGCTGCAAGCTCT 942
 DB 785 TTGACTCAAAATGCTCTTGAAGAAAGAGCCAAATGAGGCTAGTCTGCTGCAAGCTCT 844
 QY 943 AGAGTGGGAGTGTCTCCACCAACCAAGAGCTGAAGTGATGATAAGAACTTGTTT 1002
 DB 845 AGAGTGGGAGTGTCTCCACCAACCAAGAGCTGAAGTGATGATAAGAACTTGTTT 904
 QY 1003 GTGGCTACAGAGAGTCTCAACCTCAAAAAGCCAGAGTGTCTTATGATAGCCCTAC 1062
 DB 905 GTGGCTACAGAGAGTCTCAACCTCAAAAAGCCAGAGTGTCTTATGATAGCCCTAC 964
 QY 1063 AAAACTAGTATGATGAGAGGAGTCAAAAAGATCTTCTCAACCTATTAATCAAGAAAGG 1122
 DB 965 AAAACTAGTATGATGAGAGGAGTCAAAAAGATCTTCTCAACCTATTAATCAAGAAAGG 1024
 QY 1123 AATCTCTTCAAC 1133
 DB 1025 AATCTCTTCAAC 1035

RESULT 4
 WSDCUANSA
 LOCUS WSDCUANSA 2505 bp DNA linear BCT 18-JUL-1997
 DEFINITION W.succinogenes dcua and ansa genes.
 ACCESSION X89215
 VERSION X89215.1 GI:895917
 KEYWORDS ansa gene; asparaginase; C4-dicarboxylate membrane transporter; dcua gene.

SOURCE
 ORGANISM
 Wolinella succinogenes
 Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales; Helicobacteriaceae; Wolinella.

REFERENCE
 AUTHORS Lubkowski, J., Palm, G.J., Gilliland, G.L., Derst, C., Rohm, K.H. and Wlodawer, A.
 TITLE Crystal structure and amino acid sequence of Wolinella succinogenes L-asparaginase
 JOURNAL Eur. J. Biochem. 241 (1), 201-207 (1996).
 MEDLINE 8898907
 PUBMED 8898907

COMMENT
 FEATURES
 TITLE Direct Submission
 AUTHORS Roehm, K.H.
 JOURNAL Submitted (27-JUN-1995) K.H. Roehm, Institut fuer Physiologische Chemie, Philipps-Universitaet Marburg, Karl-von-Frisch-Str.1, 35033 Marburg, FRG
 Related sequences X83689.

COMMENT
 FEATURES
 source location/Qualifiers
 1..2505
 /organism="Wolinella succinogenes"
 /mol_type="genomic DNA"
 /db_xref="taxon:844"
 /clone_id="lambda EMBL3"
 1..1068
 /gene="dcua"
 <1..1068
 /gene="dcua"
 /codon_start=1
 /transl_table=11
 /product="C4-dicarboxylate membrane transporter"
 /protein_id="CAA61502.1"
 /db_xref="GI:895918"
 /db_xref="SWISS-PROT:Q34245"
 /translation="KILRSQPKHTEFLAVVYFWTLTLAGTGHFSTLPIAKVAKE
 QGVKSVLETCRGRSLATITAPSAATVPMGSGRDRGVGLIOLALICIPPTFG
 FYATATLNTDLSKDPYQERVAAGKAKGQDQYVAKPKAKLSVLLFLIGLITAV
 MATAISKNVGLIOMPVIGRDSALIMFMKLTATLAMPKIDTDVAIKASTERSGMTA
 CIGVAVMLGDTFVSNHINDIKALSGILGHPMMLITLIFASMLYSQAATAKAL

misc_feature
 1069..1188
 /note="intergenic region"
 terminator
 1079..1102
 /note="potential rho-independent terminator for dcua"
 1171..1177
 /note="putative"
 1189..2181
 /gene="ansa"
 1189..2181
 /gene="ansa"
 /EC_number="3.5.1.1"
 /EC_number="3.5.1.1"
 /codon_start=1
 /transl_table=11
 /product="asparaginase"
 /protein_id="CAA61503.1"
 /db_xref="GI:895919"
 /db_xref="SWISS-PROT:P50286"
 /translation="NAKPVTLTATGTTAGSGESSVKSYSAGAVTVKILAAVPAI
 NDLATIKGEOLISIGSOEMTGRVWLKIAKRVNELLAKQKTEAVITIGHDTMEIAF
 LNLTVKSORPVVLYGAMRPGSSADGPNLNAVNAVNAINKSTNGVYIVMDETHA
 AREATKNTTAVNAFASPNYKIGVYKVEYFTQSVPHTLASEFDISKLELRY
 DITVAHPDQTDVIVNAILOAGAKGIHAGMGNGNPPLIONALEKAKSGVVAERSR
 VSGSTTGQAEVDKDKLGFVATESLNPKARVILMLALTKTSREAIQKIFSY"
 2186..2219
 /note="potential rho-independent terminator for ansa"
 2186..2219

BASE COUNT 594 a 690 c 553 g 668 t
 ORIGIN

Query Match 90.7%; Score 1027.8; DB 1; Length 2505;
 Best Local Similarity 99.8%; Pred. No. 1.1e-284;
 Matches 1029; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 103 AGCGCTCTGTTTGTGCTAAACCCCAAGTACTATCTTACGACAGAGGACCATC 162
 DB 1174 AGCGCTCTGTTTGTGCTAAACCCCAAGTACTATCTTACGACAGAGGACCATC 1233
 QY 163 GCTGTTGGGGGAATCTAGCGTCAAGAGTACTCTGCTGAGAGTACCGTTGAT 222
 DB 1234 GCTGTTGGGGGAATCTAGCGTCAAGAGTACTCTGCTGAGAGTACCGTTGAT 1293
 QY 223 AAGCTTCTGACACCGCTCCCTGCATCAACGACCTAGCCACCATCAAGGTTGAACATC 282
 DB 1294 AAGCTTCTGACACCGCTCCCTGCATCAACGACCTAGCCACCATCAAGGTTGAACATC 1353
 QY 283 TCAAGCATTTGCTCCCAAGAGATGAGGATAGGTTAGGCTTAAGTACCAAGCGTGC 342
 DB 1354 TCAAGCATTTGCTCCCAAGAGATGAGGATAGGTTAGGCTTAAGTACCAAGCGTGC 1413
 QY 343 AATGAGCTCTGCCCCCAAGAGACCGAGCGGTGATCATCACCATGAGTACACAC 402
 DB 1414 AATGAGCTCTGCCCCCAAGAGACCGAGCGGTGATCATCACCATGAGTACACAC 1473
 QY 403 ATGAGAGACCGCTTCTCTTCTTCACTACGCTGAGAAACCAAAACCTGTCGCTT 462
 DB 1474 ATGAGAGACCGCTTCTCTTCTTCACTACGCTGAGAAACCAAAACCTGTCGCTT 1533
 QY 463 GTAGGGCCCATGCTCCAGGCTTTCATGAGTGTGAGGCCCATGAACTCTATTAAC 522
 DB 1534 GTAGGGCCCATGCTCCAGGCTTTCATGAGTGTGAGGCCCATGAACTCTATTAAC 1593
 QY 523 GCCGTGAATGAGGATCAACAAAGCCTTACTAACAAGAGTGTGATGTGTAAC 582
 DB 1594 GCCGTGAATGAGGATCAACAAAGCCTTACTAACAAGAGTGTGATGTGTAAC 1653
 QY 583 GATAGATTTACGCGCCGACAGAGAGCGCAAGCTCAACACACGAGTCAATCATTT 642
 DB 1654 GATAGATTTACGCGCCGACAGAGAGCGCAAGCTCAACACACGAGTCAATCATTT 1713
 QY 643 GCTTCGCCCAACAGATTAATGCGACAGTCTATTATGGCAAGTGGAGTATTTCAT 702
 DB 1714 GCTTCGCCCAACAGATTAATGCGACAGTCTATTATGGCAAGTGGAGTATTTCAT 1773

QY	703	CAATCGCTTGACCTCACAACCCCTTGCAAGTGAAGTTGATATAGCAAAATCGAACAATC	762
Db	1774	CAATCGCTTGACCTCACAACCCCTTGCAAGTGAAGTTGATATAGCAAAATCGAACAATC	1833
QY	763	CCGAGAGTGATATATCTTTACGCTCACCCCGAGATACATGATGTTTAACTCAATGCAGCC	822
Db	1834	CCGAGAGTGATATATCTTTACGCTCACCCCGAGATACATGATGTTTAACTCAATGCAGCC	1893
QY	823	CTTCAGGCGAGAGCCAAAGAAATCATTCATGCAGGATGGGCAATGGGAACCTTCCCT	882
Db	1894	CTTCAGGCGAGAGCCAAAGAAATCATTCATGCAGGATGGGCAATGGGAACCTTCCCT	1953
QY	883	TTGACCTCAAAATGCTCTTGAAGAAAGCAGCCAAATCAGGCGCTAGCTGCTCGAGACTCT	942
Db	1954	TTGACCTCAAAATGCTCTTGAAGAAAGCAGCCAAATCAGGCGCTAGCTGCTCGAGACTCT	2013
QY	943	AGAGTGGGCGACTGTTCCACACCCCAAGAGGCTGAGTGAATATATAAGAAACTTGGTTTT	1002
Db	2014	AGAGTGGGCGACTGTTCCACACCCCAAGAGGCTGAGTGAATATATAAGAAACTTGGTTTT	2073
QY	1003	GTGGCTAAGAGAGATCTCAACCTCAAAAGCCAGATGCTTCTTATGTTAGCCCTCAC	1062
Db	2074	GTGGCTAAGAGAGATCTCAACCTCAAAAGCCAGATGCTTCTTATGTTAGCCCTCAC	2133
QY	1063	AAACTAGTGTATAGAGAGGCGCATCCAAAGAATCTTCTCACCTATTATCCAGAAAGGG	1122
Db	2134	AAACTAGTGTATAGAGAGGCGCATCCAAAGAATCTTCTCACCTATTATCCAGAAAGGG	2193
QY	1123	AATCTCTTCAC	1133
Db	2194	AATCTCTTCAC	2204
RESULT 5	WSAJ2933		
LOCUS	WSAJ2933	3606 bp	DNA linear
DEFINITION	Mollinella succinogenes aspa, dcua genes and partial ansa gene.		
ACCESSION	AJ002933		
VERSION	AJ002933.1	GI:2644958	
KEYWORDS	ansa gene; aspa gene; aspartate ammonia-lyase; C4-dicarboxylate membrane transporter; dcua gene; L-asparaginase.		
SOURCE	Mollinella succinogenes		
ORGANISM	Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales; Helicobacteraceae; Mollinella.		
REFERENCE	1		
AUTHORS	Ullmann, R., Gross, R., Simon, J., Unden, G. and Kroger, A.		
TITLE	Transport of C(4)-dicarboxylates in Mollinella succinogenes		
JOURNAL	J. Bacteriol. 182 (20), 5757-5764 (2000)		
MEDLINE	20461222		
PUBMED	11004174		
REFERENCE	2		
AUTHORS	Ullmann, R.		
TITLE	Direct Submission		
JOURNAL	Submitted (19-NOV-1997) Ullmann R., Institut fuer Mikrobiologie, Johann-Wolfgang-Goethe Universitaet, Biozentrum N240, Marie-Curie-Strasse 9, 60439 Frankfurt, GERMANY		
COMMENT	Related sequence: X89215.		
FEATURES	Location/Qualifiers		
SOURCE	1..3606		
	/organism="Mollinella succinogenes"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:844"		
	400..1815		
	/gene="aspa"		
	400..1815		
	/gene="aspa"		
	/codon_start=1		
	/transl_table=1		
	/product="aspartate ammonia-lyase"		
	/protein_id="CAA05764.1"		
	/db_xref="GI:2644959"		
gene			
CDS			

[illegible]

Db 3595 GTAGGCCCATG 3606

RESULT 6
LOCUS AP005085/c 303450 bp DNA linear BCT 05-MAR-2003
DEFINITION Vibrio parahaemolyticus DNA, chromosome 2, complete sequence, 2/6.
ACCESSION AP005085 BA000032
VERSION AP005085.1 GI:28808465
KEYWORDS
SOURCE
ORGANISM
Vibrio parahaemolyticus
Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
Vibrionaceae; Vibrio.

REFERENCE
AUTHORS 1 Nasu, H., Iida, T., Sugahara, T., Yamachi, Y., Park, K.S., Yokoyama, K.,
Makino, K., Shinagawa, H. and Honda, T.
TITLE A filamentous phage associated with recent pandemic Vibrio
parahaemolyticus O3:K6 strains
JOURNAL J. Clin. Microbiol. 38 (6), 2156-2161 (2000)
MEDLINE 20295086
PUBMED 10834969
REFERENCE
AUTHORS 2 Makino, K., Oshima, K., Kurokawa, K., Yokoyama, K., Uda, T.,
Tegomori, K., Iijima, Y., Najima, M., Nakano, M., Yamashita, A.,
Kubota, Y., Kimura, S., Yasunaga, T., Honda, T., Shinagawa, H.,
Hattori, M. and Iida, T.
TITLE Genome sequence of Vibrio parahaemolyticus: a pathogenic mechanism
distinct from that of V. cholerae
JOURNAL Lancet 361 (9359), 743-749 (2003)
MEDLINE 22508454
PUBMED 12620739
REFERENCE
AUTHORS 3 (bases 1 to 303450)
Oshima, K., Kurokawa, K., Makino, K., Yokoyama, K., Yasunaga, T.,
Honda, T., Shinagawa, H., Hattori, M. and Iida, T.
TITLE Direct Submission
SUBMITTED (09-APR-2002) Ken Kurokawa, Osaka University, Genome
Information Research Center; 3-1, Yamadaoka, Suita, Osaka 565-0871,
Japan (E-mail: kurokawa-gen@gen.osaka-u.ac.jp,
URL: http://www.gen-info.osaka-u.ac.jp/, Tel: 81-6-6879-8365,
Fax: 81-6-6879-2047)
COMMENT
This clone was isolated from a patient presenting with acute
gastroenteritis.

FEATURES
source location/Qualifiers
1. 303450
/organism="Vibrio parahaemolyticus"
/mol_type="genomic DNA"
/strain="O3:K6"
/sub_strain="RIMD 2210633"
/db_xref="taxon:670"
/chromosome="2"
81..1574
/gene="VPA0300"
81..1574
/gene="VPA0300"
/note="similar to EMB:CA616187.1 (A1591787) percent
identity 68 in 487 aa"
/codon_start=-1
/evidence-not-experimental
/transl_table=11
/product="hypothetical signal peptide protein"
/protein_id="BAC61643.1"
/db_xref="GI:28808466"
/translation="MNLKAVPKALVTLTIASSIPSSVAYAAPNANESSLVOOONEGN
WPTONEAEKLYOELYORALSAITQOPAMNYIGMDGSESKFGKGNVLPIMKDRMD
SRTIVPTPNADVSNKYIDLEKTEGLVVAAPKVIQFTDFEORTITDVGILGDRRA
RGGLVLLPNDVGVKPYFTFTSTNNVAFPEFTVMKFGNGADSPSAVITAGOTR
IPLWDEKNIQPMKPNASGQRINMYPDTSYEMKQFNGADSPSAVITAGOTR
ASIGVKGKPNFTVEKOLLEKAVTAPKMTAORTLCRDEBNITKRDYIETAMA
GATSEYMSYLDIDORAFQYAYSSAPAMNRTMNSGSKITFTVROSKGDIENGSH
OYKHLHPANPAKLEAVITLVNVGCTMPETISQLPSKNGFDNIENKNDGSDIDYFGP
QLPNAPESNYIKITIPGRDLTAIRLYGIGIEFDQWPKPPDVLEKIK"

gene
CDS
complement(1835..2314)
/gene="VPA0301"
complement(1835..2314)
/gene="VPA0301"
/note="similar to DBJ:BA043758.1 (AP003138) percent
identity 57 in 158 aa"
/codon_start=-1
/evidence-not-experimental
/transl_table=11
/product="putative peptide methionine sulfoxide reductase"
/protein_id="BAC61644.1"
/db_xref="GI:28808467"
/translation="MMKITYAGGCLMGVQEPMRHLPGVISTEAGRANGKTDNQSEY
DGYACVTRTEPDSPQSVSETLMGYFEFTIDPYSIKNGDDGGERYRGRYSRDRHLE
IAKSFIAERYDDKLAIVELPLKYNVPSDEHODRLTFPNDYCHIFLIDLHKYRK"
2493..2882
/gene="VPA0302"
2493..2882
/gene="VPA0302"
/note="similar to GB:AAA64344.1 (U16865) percent identity
52 in 129 aa"
/codon_start=-1
/evidence-not-experimental
/transl_table=11
/product="hypothetical protein"
/protein_id="BAC61645.1"
/db_xref="GI:28808468"
/translation="MKIEHVAIWERLEKGEYKFNANVNDKYNPKHRSYFL
SFESGARLEIEMSGVTCENSHMQYGLAHFAFGSDQAVDQILKTLVEGQYRI
DGPRTYGDGYEVCVLDPDGRIELTV"
2974..3714
/gene="VPA0303"
2974..3714
/gene="VPA0303"
/note="similar to GB:AAL20622.1 (AE008775) percent
identity 54 in 243 aa"
/codon_start=-1
/evidence-not-experimental
/transl_table=11
/product="putative regulatory protein"
/protein_id="BAC61646.1"
/db_xref="GI:28808469"
/translation="MNPKNQNEILQVNDKRRVQVELSDIGSGVTIRQDINPLEQ
GYLKRVGATATALQSDIDITPLEVDIKOTLANKAADVARNEVILLEGSAANLLA
RTLAERGDAVITITPSAYIAHLIRNTSANIILIGYVQHGSEISGPIKLCIENTHFS
TAVLIGIDGFDHODGFTSRDMRADIAEAILAKRRNIVLTSSKFGQIYPSISIKTNE
ISVLTDRKAAPKSDLEOKKIGEVVLS"
complement(3733..4089)
/gene="VPA0304"
complement(3733..4089)
/gene="VPA0304"
/note="similar to GB:AAF96321.1 (AE004376) percent
identity 76 in 117 aa"
/codon_start=-1
/evidence-not-experimental
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="BAC61647.1"
/db_xref="GI:28808470"
/translation="MHOHEKINLYVEFGTPNIGATKAPPEHYFGNQFVDYGPDYAEEG
QGLDGFYSACVQSTNSGALLVIFYSSDHATIEKVAKRGGQILRIFTEPPGGRPH
FTPEPGENFAWSESH"
4319..5860
/gene="VPA0305"
4319..5860
/gene="VPA0305"
/note="similar to REF:NP_231225.1 (NC_002505) percent
identity 83 in 510 aa"
/codon_start=-1
/evidence-not-experimental
/transl_table=11
/product="catalase"
/protein_id="BAC61648.1"

/db_xref="GI:2808471"
 /translatation="MSTHMRATQMOKSEPLLITVIGASTSLQATLTRENGAPVDN
 ONSTIAGENGSVLLQDHYHLOKLOFRAPERLPEKRVYHARGTGAGEVAGDEVSDLTV
 SAPLEKPKATVPYPRRSTVTHSKSGPETLEDPFGAFATKYTEQGMNDLVGNMLPYF
 IRDSIKPEPDWYHSLKSPVTVNODPNRPFDESEPSATIMLTWYVSNLGTPTATM
 DFGVHAHAKWVINOQDVNTVYFQKSSQOGLSLRPNVTEMOGDLVYHNLIDYAAIG
 RGNFKPMQLYYKSLSPLEALSKLDNGDATVYVNLNPKRYGWTILNRLENEFLETE
 GASFAPNSLIPGIEPSEDRLLQGRLEFADYQILRLCANLPOLVYNRLPSVNNHON
 GLSNAOISGNDVNEPSPKRLNLADNDFQKVEKRLVGTVOQKALSKPDRDYQGVLY
 RSMNDQSDSLIANLAGDLINKYIDKIDKATWVSFYPRADKEGSLAEADPTNLSQVK
 NKAMX"
 5362 . 6429
 /gene="VPA0306"
 5362 . 6429
 /gene="VPA0306"
 /note="similar to GB:AAF94738.1 (AE004235) percent
 identity 63 in 152 aa"
 /codon_start=1
 /evidence-not_experimental
 /transl_table=11
 /product="AnkR protein"
 /protein_id="BAC61649.1"
 /db_xref="GI:28808472"
 /translation="MKWVEILSSVYSIGLSFLPASODLETOYOSLYELEPDAARI
 GNNSEVTFVPSQGPIDQNNNESTITALMVAAYQGNQDVTYLLDSCANNCLQDRONT
 ALMQLIKREISIKLDYQAQCSFDLNNKGLDKITPAELIYQGSFVYLQNK"
 complement(6548 . 7201)
 /gene="VPA0307"
 complement(6548 . 7201)
 /gene="VPA0307"
 /note="similar to GB:AAF42465.1 (AE002563) percent
 identity 47 in 200 aa"
 /codon_start=1
 /evidence-not_experimental
 /transl_table=11
 /product="putative pyrazinamidase/nicotinamidase"
 /protein_id="BAC61650.1"
 /db_xref="GI:28808473"
 /translation="MTVRNYQYOTRAVIDVDEPKGFSLECPDELRYAAGLDIVPELKN
 HAQRKILYRDLHPKRAMDAETPANNLEPYGLPNVYKMRHCVLQTEGLDGL
 PVLVDIEQVYKNGMPDAPRGVIFEDVDATDTGANEFLCKNKLDTYVGGVLAIDFC
 VKRSVMQADDLGKRYLVNLAATRAVLPDVDSVIVEMKEKGLFVKMADDIIVERFA"
 complement(7234 . 7959)
 /gene="VPA0308"
 complement(7234 . 7959)
 /note="similar to GB:AAF96011.1 (AE004352) percent
 identity 60 in 241 aa"
 /codon_start=1
 /evidence-not_experimental
 /transl_table=11
 /product="conserved hypothetical protein0"
 /protein_id="BAC61651.1"
 /db_xref="GI:28808474"
 /translation="MIVYDILIPRLSGCADKGLLEVLLIKRSPNRPYHGVMLPGF
 VEDDKITSEGGPADENFEARRIICERKIHTRYHRSFAFIDSGDEPDSMSINT
 HYALVDNRNVOINNAGVECOLKWFPLQALINLEETLADPHQKTIRKANQKIPASTE
 YTSVLEFLDKDEPLVADIIISAYQEGFIDISRMITIKRLLDSGVLLKPNKVASYTKGKG
 GKPANVTYLTDSVEYFPQNCNGR"

[illegible]

AUTHORS

TITLE

JOURNAL
PUBMED
REFERENCE
AUTHORS

Jin, Q., Yuan, Z. H., Xu, J. G., Wang, Y., Shen, Y., Lu, W. C., Wang, J. H.,
Liu, H., Yang, J., Yang, F., Ou, D., Zhang, X. B., Zhang, J. Y., Yang, G. W.,
Wu, H. T., Dong, J., Sun, L. L., Xue, Y., Zhao, A. L., Gao, Y. S., Zhu, J. P.,
Kan, B., Chen, S. X., Yao, Z. J., He, B. R., Chen, R. S., Ma, D. L.,
Qiang, B. Q., Wen, Y. M., Hou, Y. D., and Yu, J.
Genome sequence of *Shigella flexneri* 2a: insights into
pathogenicity through comparison with genomes of *Escherichia coli*
K12 and O157
Nucleic Acids Res. 30 (20), 4432-4441 (2002)

FEATURES
source

Location/Qualifiers
1. 10893
/organism="Shigella flexneri 2a str. 301"
/mol_type="genomic DNA"
/strain="301"
/serotype="2a"
/db_xref="taxon:198214"
188. 541
/locus_tag="SF2947"
188. 541
/locus_tag="SF2947"
/note="Residues 1 to 114 of 117 are 80 pct identical to
residues 1 to 114 of a 216 aa protein from *Klebsiella*
pneumoniae gb: AAF06334.1"
/codon_start=1
/transl_table=11
/product="orf, partial conserved hypothetical protein"
/protein_id="AA044428.1"
/db_xref="GI:24053359"
/translation="MKMSRTSLVCAISLALVSTGVNAERVAIKLHNERSHVH
OSPEELAKVEKQSLGNNVIRITPSSOMGNARETWELLQNGALDMTGSADSESPFN
558. 1151
558. 1151
/locus_tag="SF2948"
558. 1151
/locus_tag="SF2948"
/note="Residues 2 to 197 of 197 are 48 pct identical to
residues 145 to 341 of a 341 aa protein from *Bacillus*
halodurans ref: NP_241567.1"
/codon_start=1
/transl_table=11
/product="orf, partial conserved hypothetical protein"
/protein_id="AA044429.1"
/db_xref="GI:24053360"
/translation="MDSTKDGFPALSAVYAGTRSYAKPKTPKPDGLKLRVOPS
PTTIKMLMGSGPTPSPEGEVYTAQGGVDAENNVSPVMOTRIETAKESDEH
ASIDPLVISTKPMKRTPEQDQILEFAAKSEAYQOKLMEKIDADTRAQAKAMGGEI
KVVKAPRRKAVQPLDFDEKDKPKRAALLKEFDNAO"
1352. 1699
/gene="dctQ"
/locus_tag="SF2949"
1352. 1699
/gene="dctQ"
/locus_tag="SF2949"
/note="Residues 1 to 108 of 115 are 75 pct identical to
residues 22 to 129 of a 136 aa protein from *Klebsiella*
pneumoniae gb: AAF01413.1"
/codon_start=1
/transl_table=11
/product="putative small integral C4-dicarboxylate
membrane transport protein DctQ"
/protein_id="AA044430.1"
/db_xref="GI:24053361"
/translation="MTWIVGLGAAYTVGAORHLSDLFALALNKKRQGLLSTIVNL

CDS

gene

ILGFASVITGGLKLDIKTLATSOVSAMQIPMGVITLPLSGVMMFYALCFING
SIQOLKQAPQAEAS
complement(1722..2117)
/gene="insB"
/locus_tag="SF2950"
complement(1722..2117)
/gene="insB"
/locus_tag="SF2950"
/note="Residues 1 to 131 of 131 are 100 pct identical to
residues 37 to 167 of a 167 aa protein from *Shigella*
flexneri ref: NP_085393.1"
/codon_start=1
/transl_table=11
/product="orf, conserved hypothetical protein"
/protein_id="AA044431.1"
/db_xref="GI:24053362"
/translation="MIVCAEMDEHMGVGAQRQRLFYADIRRTVAHAFGEPTL
ATLERLSLTLAFEVVWMDGMPLVESRLKGLHYISKRYTORIERHNLNROHLAR
LGRSLSPKSKVEIHDKVICHYLNIRHYQ"
complement(2144..2419)
/gene="insA"
/locus_tag="SF2951"
complement(2144..2419)
/gene="insA"
/locus_tag="SF2951"
/note="Residues 1 to 91 of 91 are 98 pct identical to
residues 1 to 91 of a 91 aa protein from *Shigella flexneri*
ref: NP_085394.1"
/codon_start=1
/transl_table=11
/product="orf, conserved hypothetical protein"
/protein_id="AA044432.1"
/db_xref="GI:24053363"
/translation="NASISIRCPSCSATEGVNRNKRAGQRILCSHCRTQWOLF
YASOPGTHQKIIDMANMGVGRASARIMGVLNIVLHNLNSGRSR"
2578. 3777
/locus_tag="SF2952"
2578. 3777
/locus_tag="SF2952"
/note="Residues 1 to 395 of 399 are 61 pct identical to
residues 38 to 431 of a 435 aa protein from *Salmonella*
typhimurium L72 ref: NP_462086.1"
/codon_start=1
/transl_table=11
/product="putative integral membrane protein, possible
transporter"
/protein_id="AA044433.1"
/db_xref="GI:24053364"
/translation="MLMFPMDIAITVSQRLANGLDNPFALLAIPFTRAGTLNMSGC
DPAFSTVAVNVSCTIGLIPSNVLIVESTAGVSVAIFVAGTIPGILMLAVIY
YTFILAKRGVPLSERATFAQACAFDALPSLIVTYVGGILFATERSAIAVY
LMGSDPLILITLINIVITVIGTIPDMTPAVIFPILPLPADELGMDDPVHFGIMV
ANICIGLTPPVSAFLVGSISGVATIQHLINPLLFALVALVAMITVYIPQSLFI
POLLGLM"
complement(3878..4804)
/gene="yggM"
/locus_tag="SF2953"
complement(3878..4804)
/gene="yggM"
/locus_tag="SF2953"
/note="Residues 1 to 308 of 308 are 99 pct identical to
residues 13 to 320 of a 335 aa protein from *Escherichia*
coli K12 ref: NP_417431.1"
/codon_start=1
/transl_table=11
/product="putative alpha helix chain"
/protein_id="AA044434.1"
/db_xref="GI:24053365"
/translation="MLMTGNAMADDEPTENILKDKQKQYHILKLDLISLKNLDAK
GNQATWSREGVSSDDITVWVGOLADYELIQETWEDKRVKFSAMLTSKGPASWS
VNFISFOAASDRGVVDDIKTNKKYLIIVNSEDFNRFSGLESALNTQKNSIPALKE

CDS

gene

complement(2144..2419)
/gene="insA"
/locus_tag="SF2951"
complement(2144..2419)
/gene="insA"

CDS

gene

complement(2144..2419)
/gene="insA"
/locus_tag="SF2951"
complement(2144..2419)
/gene="insA"
/locus_tag="SF2951"
/note="Residues 1 to 395 of 399 are 61 pct identical to
residues 38 to 431 of a 435 aa protein from *Salmonella*
typhimurium L72 ref: NP_462086.1"
/codon_start=1
/transl_table=11
/product="putative integral membrane protein, possible
transporter"
/protein_id="AA044433.1"
/db_xref="GI:24053364"
/translation="MLMFPMDIAITVSQRLANGLDNPFALLAIPFTRAGTLNMSGC
DPAFSTVAVNVSCTIGLIPSNVLIVESTAGVSVAIFVAGTIPGILMLAVIY
YTFILAKRGVPLSERATFAQACAFDALPSLIVTYVGGILFATERSAIAVY
LMGSDPLILITLINIVITVIGTIPDMTPAVIFPILPLPADELGMDDPVHFGIMV
ANICIGLTPPVSAFLVGSISGVATIQHLINPLLFALVALVAMITVYIPQSLFI
POLLGLM"
complement(3878..4804)
/gene="yggM"
/locus_tag="SF2953"
complement(3878..4804)
/gene="yggM"
/locus_tag="SF2953"
/note="Residues 1 to 308 of 308 are 99 pct identical to
residues 13 to 320 of a 335 aa protein from *Escherichia*
coli K12 ref: NP_417431.1"
/codon_start=1
/transl_table=11
/product="putative alpha helix chain"
/protein_id="AA044434.1"
/db_xref="GI:24053365"
/translation="MLMTGNAMADDEPTENILKDKQKQYHILKLDLISLKNLDAK
GNQATWSREGVSSDDITVWVGOLADYELIQETWEDKRVKFSAMLTSKGPASWS
VNFISFOAASDRGVVDDIKTNKKYLIIVNSEDFNRFSGLESALNTQKNSIPALKE

CDS

gene

complement(3878..4804)
/gene="yggM"
/locus_tag="SF2953"
complement(3878..4804)
/gene="yggM"
/locus_tag="SF2953"
/note="Residues 1 to 308 of 308 are 99 pct identical to
residues 13 to 320 of a 335 aa protein from *Escherichia*
coli K12 ref: NP_417431.1"
/codon_start=1
/transl_table=11
/product="putative alpha helix chain"
/protein_id="AA044434.1"
/db_xref="GI:24053365"
/translation="MLMTGNAMADDEPTENILKDKQKQYHILKLDLISLKNLDAK
GNQATWSREGVSSDDITVWVGOLADYELIQETWEDKRVKFSAMLTSKGPASWS
VNFISFOAASDRGVVDDIKTNKKYLIIVNSEDFNRFSGLESALNTQKNSIPALKE


```

FEATURES
source
    1..292906
        /organism="Shigella flexneri 2a str. 2457T"
        /mol_type="genomic DNA"
        /strain="2457T"
        /serotype="2a"
        /db_xref="taxon:198215"
        complement(119..3487)
        /gene="recC"
        /locus_tag="S3030"
        complement(119..3487)
        /gene="recC"
        /locus_tag="S3030"
        /function="enzyme: Degradation of DNA"
        /note="residues 1 to 1122 of 1122 are 96.34 pct identical
        to residues 1 to 1122 of 1122 from Escherichia coli K-12 :
        B2822"
        /codon_start=1
        /transl_table=11
        /product="DNA helicase, ATP-dependent dsDNA/ssDNA
        exonuclease V subunit, ssDNA endonuclease"
        /protein_id="AAP18145.1"
        /db_xref="GI:30042421"
        /translation="MLRVYHNRDLVLEALMEFIVERERLDPEPEXILVQSTGMAQ
        /WLOMTLSQKFGIANIDPLPASFIMDFVAVLEIPRESAFNKQSMKMLTLPLPOL
        LEREDFTLRHYLDDSDKRLKFLQSSKADLFQYIVYRDMLAQWETGHLVEGIGE
        AQMOAPIMKALVEYHELGOPRWRNANLVORFETLESATTCPPGLSPRVEIGISA
        LPYVYLAQALQAIKHEIFHILFTNRCRYMGDIDIPAYLAKLIRORRHSFEDRELPL
        FRSENKGLPNSGDEDDVGNLSLNSKGLKRDVITYLLDSSOELDAPDYPMDNL
        LNHQSDILEENKAVAGVNIIEFSRSDNKRPLPLDSITTFHCHSPREVEYVHDR
        LAMLEAPPLTPRDIIVMAADISYSPFIQAVGSAADRYLPLAISDRRROSHVP
        LEAFISLILSDRSFSEVDYALDVPVLAARPDITEGLRYLRQWVESGIRWIDID
        DNARELELPATGQTRFGLRMLDGYAMESAQCEQSVLPYDESSGLIAELVGLAS
        LMOINWRGIAQERLEEMLPVCRMLNLFPLDPAETEAAMLIEOQMAIIIEGL
        GAAYGADVPISILDELQALRDOERISORPLAGVNICITLMPMSIIPKVCILGMD
        GVPYPLAPLIGFDLMSQPKRGDSRRDDDTLLEALISNOQKCHSYVRSIODNS
        ERPSVYQELIDITIGOSHYLPDEBALNCDESAKVAHLTLTRMPDPDQNTYQPE
        RQSYARWMLPAASQAGKASEVQPLPTLPETVLELQRFWAHPVRAFPQMRQYN
        FHTEDSEIPDTEPFLIEGLSRVQINQDLNALVDDDAERLFRRAAGDLPYGFGE
        IFWETQCEMOQLADRYIACRQPSMEIDACNQVQITGMPLQVDPGLLRMRSL
        SVQGMQLMEHLVYCASGNGESRFLFRKDEWRFPLLAEOALHYLSQLEEGREG
        MSAPLYLPESGGMKLTQYDAONAMDDSTQKATKFLQAVEGNMVVRGEGDDI
        WYORLRIQLPETHMETIVEOSRFLPLFRNQS"
        complement(3500..3850)
        /gene="ppdc"
        /locus_tag="S3031"
        complement(3500..3850)
        /gene="ppdc"
        /locus_tag="S3031"
        /function="putative enzyme: Not classified"
        /note="residues 1 to 116 of 116 are 81.03 pct identical to
        residues 1 to 107 of 107 from Escherichia coli K-12 :
        B2823"
        /codon_start=1
        /transl_table=11
        /product="prepilin peptidase dependent protein C"
        /protein_id="AAP18146.1"
        /db_xref="GI:30042422"
        /translation="MSASLKNQOGSLPEVAVVAVLWLVMLVLMVITVLSGIQR
        TSNMSIASRNOYQOLMRHGMQOTOLRAISPANMVOVNRQTSOAGCVSISVYLVSPGG
        REGMTRLHCNCO"
        complement(3835..4200)
        /gene="ygdB"
        /locus_tag="S3032"
        complement(3835..4200)
        /gene="ygdB"
        /locus_tag="S3032"
        /note="residues 16 to 121 of 121 are 96.22 pct identical
        to residues 16 to 121 of 121 from Escherichia coli K-12 :
        B2824"
        /codon_start=1
        /transl_table=11
        /product="hypothetical protein"

```

```

        /protein_id="AAP18147.1"
        /db_xref="GI:30042423"
        /translation="MLIVYGSLLIQMSQODRSPASRVSMESQSLRQALVQSLAMG
        KHAMQIDPQVDCSOTYAGTDAQVCLNLADNELLTAGIEGVSIMRTGEVTDKTIYFS
        PRGMSDFCLKCALCOLP"
        complement(4239..4802)
        /gene="ppdB"
        /locus_tag="S3033"
        complement(4239..4802)
        /gene="ppdB"
        /locus_tag="S3033"
        /function="putative enzyme: Not classified"
        /note="residues 1 to 187 of 187 are 98.39 pct identical to
        residues 1 to 187 of 187 from Escherichia coli K-12 :
        B2825"
        /codon_start=1
        /transl_table=11
        /product="prepilin peptidase dependent protein B"
        /protein_id="AAP18148.1"
        /db_xref="GI:30042424"
        /translation="MPVKEQGFSLLEVLAMAISSVLLGAEFLPALQRESLSTRK
        LALEDEIMLRVFLVAKHQRAGYCHSGSGELETIVGQDCITVQMDANGTWDRP
        VKESDQIGFRKEKHEVLETLRGATSCGKGDXTNPDAIITDFQYVRQDVSGSPVL
        TVNMRASSEPGTVVDASTYGFNL"
        complement(4793..5263)
        /gene="ppdA"
        /locus_tag="S3034"
        /note="point mutation"
        /psuedo
        complement(5447..6241)
        /gene="thyA"
        /locus_tag="S3035"
        complement(5447..6241)
        /gene="thyA"
        /locus_tag="S3035"
        /function="enzyme: 2-Deoxyribonucleotide metabolism"
        /note="residues 1 to 264 of 264 are 99.62 pct identical to
        residues 1 to 264 of 264 from Escherichia coli K-12 :
        B2827"
        /codon_start=1
        /transl_table=11
        /product="thymidylate synthetase"
        /protein_id="AAP18149.1"
        /db_xref="GI:30042425"
        /translation="MKQYLELMQKVDGKNDKRGCTGLSTFGHMRFLQDGFPL
        VTKRCHLRSTIHELMLFLOGDNTIATLHNNNTYDPERADENQDLPVYQKQRANP
        TPDGRHIDQITVNLNDKNDPSRRTIVSAMNVEGLMKALAPCHAFQFVADGKLS
        QLYORSQDVFGLPENIASYALVHMAQCCDEVDPEWFGDTHLYSNHMDQTHL
        QLSREPRPLKLIIRKRESIFVRFEDFEIEGYDPHPYKAPVAI"
        complement(6248..7123)
        /gene="lgt"
        /locus_tag="S3036"
        complement(6248..7123)
        /gene="lgt"
        /locus_tag="S3036"
        /function="enzyme: Macromolecule synthesis, modification:
        phospholipids"
        /note="residues 1 to 291 of 291 are 95.87 pct identical to
        residues 1 to 291 of 291 from Escherichia coli K-12 :
        B2828"
        /codon_start=1
        /transl_table=11
        /product="phosphatidylglycerol-prolipoprotein
        diacylglyceryl transferase; a major membrane phospholipid"
        /protein_id="AAP18150.1"
        /db_xref="GI:30042426"
        /translation="MTSSYLHPEPFDVIFSIGFVALHMYGLMVGVFIAMWLATR
        ANRGSGWTKNEVNLVAGFLVGLGRIGVLYVNPQPMADPLFLRWDSGMSF
        HGLGIVVMILIPARTKSEFQVSEFLALIPFGAGACRLGNFINGELGVLFILNLYI
        PFMALPFGSRTEIDILQTNPMOWSIFDQTVGVRHPSQLEYELLEGVLFILNLYI
        RKRPRGAVSGPLIGAGARILYVEFRPDAPOTGAMVQYISNGQLISIMYIAGVI
        MMWAVRRSPQOHVS"
        complement(7274..9520)

```

gene

/gene="ptsP"
 /locus_tag="S3037"
 complement(7274. .9520)
 /gene="ptsP"
 /locus_tag="S3037"
 /function="enzyme; Transport of small molecules; Other."
 /note="residues 1 to 736 of 748 are 99.59 pct identical to
 residues 1 to 736 of 748 from Escherichia coli K-12 :
 B829"
 /codon_start=1
 /transl_table=1
 /product="PTS system, enzyme I, transcriptional regulator
 (with NPR and NMR proteins)"
 /protein_id="AAP8151.1"
 /db_xref="GI:3004247"
 /translation="MTRLRLEIYKVASAPRLNEILNLTVDICLAMDTEGVSYLAD
 HDRCYVLTMTGKIGKRGFLVLTADSEIVGIVSRLEPILNLDADQKHPEFKTIPSV
 ALKEFYRFLGQITIORQLGLVVOORELQYDESEFVLTATQAAITLQASQGLT
 ALQFYKOTRIRALPAVAVIAEGMDATPLMEQVQASLPDLERELTGALSEEL
 DANEFRRYKRFPAAGAKETRALEFDLXSHLSIDTRLRRLPAEDNGSVAMAKVTYV
 EKFAPQPLASDNDLKRADLALQORLFLHDDNQGCPNAMPFRFLVADLSATTYV
 LAELPQRVGVVVNDCAANSHAIWVRALGFLPYWGAQIDQSVLHRTLLVYDEI RKEE
 LLVPEPVEYVQRLISEELSESRLEADVNLPAQKGEIKYMLANLQSPHEKHE
 LSGCIDGLGATREITPEMLQSGSEEEVOAQCGLQMFNKRPTYLTLLDYGADKDL
 PYMLPSEEDGICGGRIRITLDDPELFILOVRAMLRANATGNTLILPPTSLDEVDQ
 EARRLIEPAGREVEDEMGYLTIPRPGIMLPEVPMPSYFMPLGKLKIVDLSVGTNDQ
 TLLAVDRNNRVRVATIDSHLPAMLRALMARAELTIDTLRLCGENAGDPMCVAILL

CDS

Query Match	26.5%;	Score 300;	DB 1;	Length 292906;
Best Local Similarity	57.1%;	Pred. No. 1.6e-74;		
Matches 568;	Conservative	0;	Mismatches 420;	Indels 6;
				Gaps 1;

QY	119	TTGGCTAAACCCCAAGTACTCTTCCTAGCGACAGGAGGACACATCGCTGGTTTCGGGGGAAT	178
QY	120985	TTGGCATATTCCCAATATACACATTTTATAGCACCCGGCGGGACATTGCCGGTGGTGGTACT	120926
QY	179	CTAGCGTCAAGAGTATGACTACTCTGCTGGAGCAGTCAACGTTTGATTAAGCTTCTTGACCGG	238
Db	120925	CGGCACCAAAATCTACTACTACACAGCGGGTAAAGTTGGCGATGAAAAATCTGTTAATGGG	120866
QY	239	TCCCTGCCATCAACGACTAGCCACCATCAAGGGTGAACGATCTCAAGACTTGGTCCC	298
Db	120865	TGCGCGACATTAAGAAACATTGGGAACGTTAAAGCGAGACAGTGTGAATATCGGGTCCC	120806
QY	299	AAGAGATGACGGGTGAAGGTGGTGGCTTAATAGCCAAACCGGTCAATAGACTCTCCGCCCC	358
Db	120805	AGGACATGAAACGATATGTCTGGCTGCAACTGCGGAAAAAATTAA-----CGCGACT	120752
QY	359	AAAAAGACCGAAGCCGATCATCACCATGCAACTGACACCATGGAAGAGACCGCTT	418
Db	120751	GGCATTAACCGACGGCTTGGCTATTTACCAGCGGTACGACATGGAAGAAATCTCTT	120692
QY	419	TCTTCTCTCAACTCAGCGGTGAAAAACCCAAACCTGTCTGGCTTGTAGGGCCATGGGTGC	478
Db	120691	ACTTCTCCACACTGACGGTGAATCGCAAAACCGGTGTATGATGGCGCAATGGGCC	120632
QY	479	CAGGCTCTTCATGAGTGTGATGGCCCCCATGAATCTGTATTAAGCGGTGATATGTACGA	538
Db	120631	CGTCCAGTGCATGAGAGCGCAGACGGTCACTTCAACTGTATTAAGCGGTTATGTACGCCAG	120572
QY	539	TCAACAAAGCTCTTACTTAACAAAGAGTGTGATTTGTATGACAGATGAGATTACGCCG	598
Db	120571	CTGATTAAGAGCTCCGCTTAATCTGTGGGGTGGTGGTGTATAGACACCCCTACTGTGAGG	120512
QY	599	CCAGAGAAGGACCAAGATCACAACACCGACGGCATGAATGCATTGGCTTCCGCCAACAG	658
Db	120511	GTCGCGATGTCAACCAAAACCAACACCGACCGATGACACTTCAAGTCTGTTAACTAAG	120455
QY	659	GTAAAAATCCGACAGCTATATTATGGCAAAAGTCAGATTTTCACTCAATCCGTTGACCTC	718
Db	120451	GTCTCTGGGATATCAATTCACAACAGGTAAATGATTGATTACCAAGATACCCCGCAGTAAGC	120392
QY	719	ACACCTTGGCAAGTGAATTTGATATTAGCAAAATGGAAGACTCCCCAGATCGATATTC	778

Db 120391 ACACGAGCATAGCCATTCGATGTCGTCTGAAGCTGAATGAGCTGCCGAAAGTGGGCTCG 120332

QY 779 TTTACGCTACCCCGATGATGATGTTTATGTCATGCAAGCCCTTCAGGAGAGACCA 838

Db 120331 TTTATACATACGCTAACGCGATCCGATCTCCGGCTAAAGCACTGTAAGATCCGGGCTATG 120272

QY 839 AAGGATATCATCCATGCGAGGCTGGGCAATGGGAAACCTTTCCCTTGACTCAAAATGCTC 898

Db 120271 ATGGCATCGTTAGCGCTGGTGTGGTAATGTAATGTAACCTGTATTAATCCGTGTTCGACACCC 120212

QY 899 TTGAAAAAGCAGCCAAATCAGCGTAGTCGTCGCTCGAAGCTCTAGAGTGGCAGTGT 958

Db 120211 TGGCAACCGCGGGAAAAAAGCGAGCTAGCTGATGTCGTCTTCGCCGTACCGAGGGGTG 120155

QY 959 CCACGCCCAAGAGGCTGGAAGTGATGATGAATAAACTGGTTGTGCTACAGAGATC 1018

Db 120151 CTACACCTCAGGATGCTGGAAGTGATGATGCAAAATACGGCTGTGCTGCCCTCTGGCAAGC 120092

QY 1019 TCAACCTCAAAAAAGCCAGAGTCTTCTTATGTTAGCCCTACCAAAACTGATGATAG 1078

Db 120091 TGAACCGGCAAAAAAGCGCGGCTCTGCTCAGCTGGCTCTGACCAACCAAGATCGC 120033

QY 1079 AGGCGATCAAAAGATCTCTCCACTATTAACT 1112

Db 120031 AGCAGATCCAGCAATTTTCAATCACTACTAATC 119998

RESULT			
9			
AE005526/c			
LOCUS			
DESCRIPTOR			
	10749 bp	DNA	linear
AE005526			BCT 21-MAR-2001
Prochlorococcus cci	C157:H7 F0R33 genome.		contig 3 of 3, section 145

ACCESSION	AE005526	AE005174
VERSION	AE005526.1	GI:12517505

SOURCE	Escherichia coli O157:H7	EDL933
ORGANISM	Escherichia coli O157:H7	EDL933
Phylum	Bacteria: Proteobacteria	Gamma proteobacteria
Class	Bacteria: Gammaproteobacteria	Enterobacteriales
Order	Bacteria: Gammaproteobacteria	Enterobacteriales
Family	Bacteria: Gammaproteobacteria	Enterobacteriaceae
Genus	Bacteria: Gammaproteobacteria	Escherichia
Species	Bacteria: Gammaproteobacteria	Escherichia coli

REFERENCE
AUTHORS
1 (bases 1 to 10749)
Perna, N. T., Plunkett, G. III, Burland, V., Mau, B., Glasner, J. D.,

Grotbeck, E.J., Davis, N.W., Lim, A., Dimalanta, E., Potamoudis, K., Anodara, I., Mantharaman, T., S. Lin, J., Yen, G., Schwartz, D.C.,

TITLE genome sequence of enterohaemorrhagic *Escherichia coli* O157:H/
JOURNAL Nature 409 (6819), 529-533 (2001)

PUBLISHED
 11/20/53
 REFERENCE
 2 (bases 1 to 10749)
 AUTHORS
 Berra, N. T., Plunkett, G., III, Burland, V., Mau, B., Glasner, J. D.,

Postaf, G., Hackett, J., Klink, S., Boutin, A., Shao, Y., Miller, L., Grotbeck, E. J., Davis, N. W., Lim, A., Dimalanta, E., Potamoudis, K.

TITLE Direct Submission
JOURNAL Submitted (22-OCT-2000) Laboratory of Genetics, University of

FEATURES	location/Qualifiers
source	1. .10749

```

/strain="EDL933"
/serotype="O157:H7"

```

gene complement(89, .1135) /note="enteronemorrhagic"

```
CDS complement(89, .1135)
```

```

/gene="ansb"
/function="enzyme; Degradation of small molecules: Amino acids"
/note="Residues 1 to 348 of 348 are 98.85 pct identical to residues 1 to 348 of 348 from Escherichia coli K-12 Strain MG1655: B2957"
/codon_start=1
/transl_table=11
/product="periplasmic L-asparaginase II"
/protein_id="AAG58088.1"
/db_xref="GI:12517506"
/translation="MEFFKRLALVWGFSAALALPNTITLTAATGTTAGGDSATK
SNITGKGVENLVNAVLPOLKDIANVKGEOVYVIGSDMDVWLTLAKIINTDDKT
DGEVTHGDTMEETAYFLDLIVKCDKPVVWVGARPESTMSADGPFNLVNAVVAAD
KASANRGVLYVMDVLDGRDVTNTITVDVATFKSVNYGPLYGINKGIDVQRTARK
HTSDTPDVSATINELPKVGIYVNYANASDLPAKALVDAGDGIIVAGVNGMLYISVE
DLTAAKNGTAVVSSRVPGATVQDAEVDAAKKGFTASGTLNPKARVLLQALTLQ
TKDPOIQIIFNOV"
/complement(1311..2030)
/gene="y99N"
/note="synonym: 24303"
/complement(1311..2030)
/gene="y99N"
/function="orf; Unknown function"
/note="Residues 1 to 239 of 239 are 99.58 pct identical to residues 1 to 239 of 239 from Escherichia coli K-12 Strain MG1655: B2958"
/codon_start=1
/transl_table=11
/product="orf; hypothetical protein"
/protein_id="AAG58089.1"
/db_xref="GI:12517507"
/translation="MMKMLAALSTAMAHADYOCSTPRDDVIVSPOTVOYKE
NGNVTTPDGVMTNGKQYSLNAAQREKADYQELRSTLPWIDEGASREKARIAL
DKIITVQEGESSKMSRLTKLDAQLKQDMNRIETRSGLTFHYKALIDVQVXEQOLV
NQAGGILQDSINEMGAKAVLGSGNPLQNVLSLGGSSIQTEMKQKQDFQFGK
DVCSTVYTLDESRAALVGNLK"
/complement(2214..2570)
/gene="y99L"
/note="synonym: 24304"
/complement(2214..2570)
/gene="y99L"
/function="orf; Unknown function"
/note="Residues 1 to 118 of 118 are 99.15 pct identical to residues 1 to 118 of 118 from Escherichia coli K-12 Strain MG1655: B2959"
/codon_start=1
/transl_table=11
/product="orf; hypothetical protein"
/protein_id="AAG58090.1"
/db_xref="GI:12517508"
/translation="MGLNVRGEETMAKRSRLRKMHIDFQELGFSVAPRPEGIS
EEOIDKVDDEINVEIPEPKLADGSSYLAWEGLICQELGKTEEHQALVRKWLERR
NIDBVRTSELFQVWMD"
/complement(2540..3259)
/gene="y99H"
/note="synonym: 24305"
/complement(2540..3259)
/gene="y99H"
/function="orf; Unknown function"
/note="Residues 1 to 239 of 239 are 99.58 pct identical to residues 1 to 239 of 239 from Escherichia coli K-12 Strain MG1655: B2960"
/codon_start=1
/transl_table=11
/product="orf; hypothetical protein"
/protein_id="AAG58091.1"
/db_xref="GI:12517509"
/translation="MRNDVISPEDENGRLRIRSEFVRQGRGLKGOEHALENPMY
MGVEFSEMDLDFALGRLAPVTLGFGGASLVAMAKRPDODFLGTEVHSPPGVA
CLASAEGLSNLRVWCMDAVELHKMIDNNSLRMOLPRPPMHRKHRRKRIYQVP
PAELVSKTQIGGVFMATDMPYAAHMLEVMSIDGIKNLSESDYIVPPAPRPVTK
PDRGRHRLGHGVMDLMFERVK"

```

```

gene
3420..4472
/gene="muly"
/note="synonym: 24306"
3420..4472
/gene="muly"
/function="enzyme; Macromolecule synthesis, modification: DNA - replication, repair, restr./modification"
/note="Residues 1 to 350 of 350 are 99.42 pct identical to residues 1 to 350 of 350 from Escherichia coli K-12 Strain MG1655: B2961"
/codon_start=1
/transl_table=11
/product="adenine glycosylase; G.C. --> T.A. transversions"
/protein_id="AAG58092.1"
/db_xref="GI:12517510"
/translation="WASQFSAQVLDWYDKYGRKTLPMQIDKTPKYVWLSEVLMQTO
VATVPIFEREMARPPVYTDLANAPLDEVLHMTGLGYARARNLHKAQVATLHGK
KEPTEFEVVALPGVGRSTAGAILSLGKHPPLIDGVNRYRLACVAVSGMPKKEY
ENKLTSLSEQVTPAVGERVENQAMDIGMTCITRSKPCSCICPONGCITAAVNSSL
YPGKKPKOTLPERGTGYFILLQHEDEVLLAQRPSCTMGICYCFPPADESIRQMLAO
ROIADNLQGLAFPHHTSHRLDILVPMMLPVSSFTGCMDSGNALMYNLADPPSVGLA
APVERLQQLRTGAPV"
4500..4775
/gene="y99X"
/note="synonym: 24307"
4500..4775
/gene="y99X"
/function="orf; Unknown function"
/note="Residues 1 to 91 of 91 are 100.00 pct identical to residues 1 to 91 of 91 from Escherichia coli K-12 Strain MG1655: B2962"
/codon_start=1
/transl_table=11
/product="orf; hypothetical protein"
/protein_id="AAG58093.1"
/db_xref="GI:12517511"
/translation="MSRTICTFLQREAGODPOLYPBELKRTYNEISKEAWAQMOH
KQYMLINEKRLNMNAEHRKLLQEMVNFLEGREVNHIEGYTPEDRK"
4837..5919
/gene="m1tc"
/note="synonym: 24308"
4837..5919
/gene="m1tc"
/function="putative enzyme; Cell envelop. Murein sacculus, peptidoglycan"
/note="Residues 1 to 360 of 360 are 99.72 pct identical to residues 1 to 360 of 360 from Escherichia coli K-12 Strain MG1655: B2963"
/codon_start=1
/transl_table=11
/product="membrane-bound lytic murein transglycosylase C"
/protein_id="AAG58094.1"
/db_xref="GI:12517512"
/translation="MMKRYLALIALPILISCTTKKGTYNEMAVKDTNGFDLMQ
FAHNIEINMGFEKVYVAGPKDYVYKTPQYOTRSIINDDGTITETIAGTEPAHNRK
AIIKTLMGDDPSVDLYSDVDOTTIKKEPFLISQVVDNGQPLRMGSRASNSADYLL
KNRLKRSNGLRITLYSVTINMVPNHLDRKRAHKYIGVROASRKGVDESILIMQTE
SSENPYAVSSDALGLMQVVOHAGKDVFSQGSQTSRSLFDPASINDTGTAYIA
MLNNVYLGIDNPTSRRYAVITAYNGAGGSVLRFVSNDKIOAANIINTMPPGVYQTL
TTRHPASRSRRYLYKVTAAQSYRRR"
6073..7377
/gene="nupg"
/note="synonym: 24309"
6073..7377
/gene="nupg"
/function="transport; Transport of small molecules: Nucleosides, purines, pyrimidines"
/note="Residues 1 to 434 of 434 are 99.76 pct identical to residues 1 to 434 of 434 from Escherichia coli K-12 Strain MG1655: B2964"
/codon_start=1
/transl_table=11
/product="transport of nucleosides, permease protein"

```

```

/ protein_id="AAG58095.1"
/db_xref="GI:12517513"
/translation="MCFKHSASRCEETINMNLKILITLFLQCLMGXLTLLGSLY
MFTYLFEDASIGAVTSLGIAVMPPLGINTISYRLONAGMDIVTDPPIRINGTGFI
AOVTEBAMLVILINSFAMPTGLINTISYRLONAGMDIVTDPPIRINGTGFI
MAMVVSLSGFEHMOVLIGALSAIVLFTLPHIPAKQAQMSITLIGLDAF
ALFKRMAIFEFISMLGAELOITNNFGTFLHSDKDFMFSFVQHASLMSIS
OISEPLILIPFLSRGILKYNMISIVMILFLFAVAGDPFGLVLTSLMIV
GCAFDFFNIGSVFPEKEVSPAIRASAGOEILGMLTNGFCILGIVSGKVEVYVYONG
ITDMOTWILIFAGISVTLARAFYAMERKIVRPITGIVYSH"
complement(7427..9622)

Query Match      26.2%; Score 296.8; DB 1; Length 10749;
Best local similarity 56.9%; Pred. No. 8.1e-74;
Matches 566; Conservative 0; Mismatches 422; Indels 6; Gaps 1;

119 TGGCTAAACCCAGTAGCTATCTCTAGCCACAGAGGACACATCGTGTTCGGGGCAAT 178
1074 TGGCATATCCCAATATCCATTTTATGACACCGCGGGACCATTCGCGGTGTGTGCTACT 1015
179 CTAGGCTCAAGTAGTACTCTCTCTGAGCAGTACCGCTGATTAAGCTTCTGCAAGCCG 238
1014 CCGCACCAATCTTACTACACAGGGGTAAAGTTGGCGTAAATCTGTTAATGCGG 955
239 TCCCTGCCATCAAGCAGCTAGCCACCATCAAGGGTGAACAGATCTCAAGCAATGGCTCC 298
954 TCCCGCACTGAAGACATTTGGCAAGTTAAAGGCGACAGTAGTGAATTCGGCTCC 895
299 AAGACATGAGGGGTAGAGTGGCTTAAACTAGCCACGCTGTCANTAGATCTCTGCGCC 358
894 AAGACATGAGCATGATGCTGTGCTGACACTGGCGAAAAAATTA-----CACCGACT 841
359 AAAAAAGACCGAAGCGGTGATCATCAATCCATGAACTGACACCATGGAAGAGACGGCTT 418
840 GCGATTAACCGACGGCTTCTGCTATCTACAGGTAACGACGATGAGGAAGAAACGGCTT 781
419 TCTTCTCAACCTCCAGCGGTGAAGAAAGCAAAACCTGTCTGCTGTAGCGGCAATCGCTC 478
780 ACTTCTCGACCTGAGCGGTGAAGTGGACAAACCGGTGTGATGTCGGCGCAATCGCTC 721
479 CAGGCTTTCATGAGTGTGATGAGCCCATGATCTCTAATACGCGCGTGAATGTGCGCA 538
720 CGTCCAGCTATGAGGCGAGAGCGTCAATCAACCTGATTAACGGGAGTGTGACCGCAG 661
539 TCAACAAGCCTTACTAACAAGAGGATGATGATGATGATGATGATGATGATGATGATG 598
660 CTGATTAAGCGCTCGCTAATCGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 601
599 CAGAGAGAGCGACCAAGCTCAACACCGCGCATGATGATGATGATGATGATGATGATGATG 658
600 GTGCGCATGTGACCAAAACCAACACCGCGCATGATGATGATGATGATGATGATGATGATG 541
659 GTAAATTCGACAGTCTATTAATGGAAGAGTGCATGATGATGATGATGATGATGATGATG 718
540 GTCCGTGGGATATTCACACCGTGAATGATGATGATGATGATGATGATGATGATGATGATG 481
719 ACACCTCTGCAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 778
480 ACACCAAGGACAGCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 421
779 TTTACGCTACCCCGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 838
420 TTTAATACCTAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 361
839 AAGGATCATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 898
360 ATGGCATGTGAGCGCTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 301
899 TTTAAAAAGCAGCCCAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 958
300 TGGCGACCGCGCGCAAAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 241
959 CCACGCCCAAGAGGCTGAAGTGAATGAAGAACTTGTGTTGTGCTACAGAGAGTCT 1018

```

```

RESULT 10
AP002563/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS
Makino, K., Yokoyama, K., Kubota, Y., Yutsudo, C. H., Kimura, S.,
Kurokawa, K., Ishii, K., Hattori, M., Tatsuno, I., Abe, H., Iida, T.,
Yamamoto, K., Ohnishi, M., Hayashi, T., Yasunaga, T., Honda, T.,
Sasakawa, C. and Shinagawa, H.
Complete nucleotide sequence of the prophage VT2-Sakai carrying the
verotoxin 2 genes of the enterohemorrhagic Escherichia coli O157:H7
derived from the Sakai outbreak
Genes Genet. Syst. 74 (5), 227-239 (1999)

JOURNAL
MEDLINE
PUBMED
10734605
20198780
10734605

REFERENCE
AUTHORS
Ohnishi, M., Mura, T., Nakayama, K., Kuhara, S., Hattori, M.,
Kurokawa, K., Yasunaga, T., Yokoyama, K., Makino, K., Shinagawa, H. and
Hayashi, T.
Comparative analysis of the whole set of rRNA operons between an
enterohemorrhagic Escherichia coli O157:H7 Sakai strain and an
Escherichia coli K-12 strain MG1655
Syst. Appl. Microbiol. 23 (3), 315-324 (2000)

JOURNAL
MEDLINE
PUBMED
11108008

REFERENCE
AUTHORS
Yokoyama, K., Makino, K., Kubota, Y., Matanabe, M., Kimura, S.,
Yutsudo, C. H., Kurokawa, K., Ishii, K., Hattori, M., Abe, H., Iida, T.,
Yamamoto, K., Hayashi, T., Yasunaga, T., Honda, T., Sasakawa, C. and
Shinagawa, H.
Complete nucleotide sequence of the prophage VT1-Sakai carrying the
Shiga toxin 1 genes of the enterohemorrhagic Escherichia coli
O157:H7 strain derived from the Sakai outbreak
Gene 258 (1-2), 127-139 (2000)

JOURNAL
MEDLINE
PUBMED
20564182
11111050

REFERENCE
AUTHORS
Hayashi, T., Makino, K., Ohnishi, M., Kurokawa, K., Ishii, K.,
Yokoyama, K., Han, C.-G., Ohtsubo, E., Nakayama, K., Mura, T.,
Tanaka, M., Tobe, T., Iida, T., Takami, H., Honda, T., Sasakawa, C.,
Ogasawara, N., Yasunaga, T., Kuhara, S., Shiba, T., Hattori, M. and
Shinagawa, H.
Complete genome sequence of enterohemorrhagic Escherichia coli
O157:H7 and genomic comparison with a laboratory strain K-12
DNA Res. 8 (1), 11-22 (2001)

JOURNAL
MEDLINE
PUBMED
11258796

REFERENCE
AUTHORS
Ohnishi, M., Kurokawa, K., Makino, K., Yasunaga, T., Shinagawa, H. and
Hayashi, T.
Direct Submission
Submitted (26-JUN-2000) Ken Kurokawa, Osaka University, Genome
Information Research Center, 3-1, Yamadaoka, Suita, Osaka 565-0871,
Japan (E-mail: ken@gen-info.osaka-u.ac.jp)

```

URL: <http://www.gen-info.osaka-u.ac.jp/>, Tel: 81-6-6879-8365,
Fax: 81-6-6879-2047)
genome project.

FEATURES

Location/Qualifiers

1..26658

/organism="Escherichia coli O157:H7"

/mol_type="genomic DNA"

/strain="0157:H7"

/sub_strain="RIMD 0509952"

/db_xref="taxon:83334"

/gene="Ecs3650"

complement(244..693)

/gene="Ecs3650"

complement(244..693)

/gene="Ecs3650"

/note="similar to B2790_ECOLI g11789154 percent identity 100 in 149 aa (Conserved in E.coli K-12)"

/codon_start=1

/evidence=not_experimental

/transl_table=11

/product="hypothetical protein"

/protein_id="BAB37073.1"

/db_xref="GI:13363122"

/translation="MAEIGIFGTMVGNLSVAEAEALITAGHKATVEDEPESDM

LPYQDYVLYVSTTQGGDLPDIPVLFQIGKIDSLGEPNLRGVATLADSSVAVNCN

GGKQFDALLQEQSAQVGEMLIDASENPEPESNPWEQNGTLLS"

complement(711..1493)

/gene="Ecs3651"

complement(711..1493)

/gene="Ecs3651"

/note="similar to YQCB_ECOLI g11789155 percent identity 99 in 260 aa (Conserved in E.coli K-12)"

/codon_start=1

/evidence=not_experimental

/transl_table=11

/product="hypothetical protein"

/protein_id="BAB37074.1"

/db_xref="GI:13363123"

/translation="MEILYODEMLVAVKSGMVLVHRSWLDREKVVVQVVRDQIS

QHVFAHRLDPTSGVLLMGLSSEAGRLAQQEHOIOKRYHAIVRGMLPEAVTQY

PLVEBLDITADFAREDKPGPAVTHYRGATVEMPVATGRYPTTRVGVLEKERTG

KHQRRLHAIHRPIIGDSKHDLRONSGAHEGLOIRIMAAASLSLTHPFTSEPLT

HAQGLDITWMAISQFGRGLPENREPEFSAPSGDGERSS"

complement(1493..1822)

/gene="Ecs3652"

complement(1493..1822)

/gene="Ecs3652"

/note="similar to B2792_ECOLI g11789156 percent identity 98 in 109 aa (Conserved in E.coli K-12)"

/codon_start=1

/evidence=not_experimental

/transl_table=11

/product="hypothetical protein"

/protein_id="BAB37075.1"

/db_xref="GI:13363124"

/translation="MTTHDRVRLQALALREHOMRDEPQHPQNSTQPFMDT

MEPEMLQWVLIPEMDLNNQDLPFAFVAVAYEALATDHPORALLIIELEKIDA

LEADNAS"

complement(1865..2233)

/gene="Ecs3653"

complement(1865..2233)

/gene="Ecs3653"

/product="Csrb RNA"

/note="inhibitor of Csrb"

complement(2444..2989)

/gene="Ecs3653"

complement(2444..2989)

/gene="Ecs3653"

/note="similar to SYD_ECOLI g11789157 percent identity 98 in 181 aa (Conserved in E.coli K-12)"

/codon_start=1

/evidence=not_experimental

/transl_table=11

/product="Syd protein"

/protein_id="BAB37076.1"

/db_xref="GI:13363125"

/translation="MODLIRQALDKFTARYCDAMHEHNSWPLSEELYGVSPSPCIIST

TRDAYVWQPOPTEGEQNVANAVAFVIVQPTIHFTYTGAGMDHQAQGIKILLLO

THSEDDFRKVOEMLIGHVYOKRKLPLPTLALTEEELEVISVGNLSEVKEKELTGT

KKRTHLAPLAEFLNOLKPLL"

3057..3905

/gene="Ecs3654"

3057..3905

/gene="Ecs3654"

/note="similar to YQCD_ECOLI g11789158 percent identity 98 in 282 aa (Conserved in E.coli K-12)"

/codon_start=1

/evidence=not_experimental

/transl_table=11

/product="hypothetical protein"

/protein_id="BAB37077.1"

/db_xref="GI:13363126"

/translation="WSTYANHQALAGLTGKSTDYRDYDASLLQGVRLNRPDLGL

KADNLPFGQIDWTLVLYELSWLNAKGLPQAVAGHVELDYTVNLIESKSPKLYNSFQ

TRFNWDEVRQTLERDLSTCAQGEVSVALYRLDEEGQIGFENGCTIDODITDYN

ETFDYDLNATSGEKVVEETLYSHLKNCLITHQPMGSIQIOYRGRQIDREKLLNY

LVSEFHNHEFHQCVERTFNLLFCQPEKLSVARTTRRGGLDINFRNSDVPST

TRLYVQ"

4017..5381

/gene="Ecs3655"

4017..5381

/gene="Ecs3655"

/note="similar to YQDH_ECOLI g11789159 percent identity 99 in 454 aa (Conserved in E.coli K-12)"

/codon_start=1

/evidence=not_experimental

/transl_table=11

/product="hypothetical protein"

/protein_id="BAB37078.1"

/db_xref="GI:13363127"

/translation="MTHISPLGSMMDLSOLEYDMLKRTASSDLYOLFNCISLAVLNS

GLSDNSKELSEFSEFNDINDNSVGRVKEELINPEPEAFVDRITRALQANLFAVL

DILEFYQKINTVTRFPLNDNSVGRVKEELINPEPEAFVDRITRALQANLFAVL

ENELVYARRVGNGLGRLNINCCGCGAMEAPKGAAGAAGQRYKDSRTFGMTEPS

IIAEPVLRVNLIIIMPDIKRLAEVRIAHGIIIFPGVGTAEELTLGILKNMFA

NKQOVLPLITGKESADYFRVLDDEFVHTLGENARHYRIIIDAAYAROMKRSMP

LVKENRDGTGDAYSFMSMRAPDLQMPPEPSHNNANLKVDPQVEVLALDRAR

SGTAVAGVKEVGIRAIERFGPYKINDKELMRMDLLGFGVAQHRMKLPGSAYTPCY

EICT"

5938..7227

/gene="Ecs3656"

5938..7227

/gene="Ecs3656"

/note="similar to SDAC_ECOLI g11789160 percent identity 100 in 429 aa (Conserved in E.coli K-12)"

/codon_start=1

/evidence=not_experimental

/transl_table=11

/product="probable serine transporter"

/protein_id="BAB37079.1"

/db_xref="GI:13363128"

/translation="METTOTSTIASKDSRSAMRKTDTPWMLGLYGAIGAVLPLPIN

AGVGMIPLIIMAILAFPMTEFAHRLGRLRVLSGNGEDTTEVEEHEFGCGAKLIT

ILYFAIYPIILVSVALTNTVESPMHSQDLGTPPRPRLISLLIIVGMITVREGDM

IVKASILVPEPVGLMILALYLIPONGALLETSLDTSATGNGMLMTLMLAPM

VSEFNHSPILISFPAKREEGDAEBCSKILAFAHIMVLTVMFVFCVSLTPA

DLAAKRONISLISLANHFNAPVYAMAPILAIATIKSFLGHYLGARFGNGMVK

SLRGKASIEINKLRITLMLVTVIVATINSIIGMETLGLPIIAMILFLMPRY

AIQKVPARKRSYSGHSNFFVVMGLIAISAFYSLSF"

7285..8652

/gene="Ecs3657"

7285..8652

/gene="Ecs3657"

Query Match 26.2% Score 296.8; DB 1; Length 26658;
Best Local Similarity 56.9%; Fred. No. 1.4e-73;
Matches 566; Conservative 0; Mismatches 422; Indels 6; Gaps 1;

[illegible]

repeat_region

CDS

/locus_tag="c5622"
 /complement(join(4833..5855,5857..5931))
 /gene="prfb"
 /locus_tag="c5622"
 /function="factor; Proteins - translation and modification"
 /note="Rf-2: Escherichia coli K-12 ortholog: b2891; Escherichia coli O157:H7 ortholog: z4229; ribosomal slippage: an in-frame uga termination codon is located within the prf sequence, and a naturally occurring +1 frameshift is required for synthesis of Rf-2"
 /codon_start=1
 /transl_table=11
 /product="peptide chain release factor 2"
 /protein_id="AAN81919.1"
 /db_xref="GI:26109714"
 /translation="MFEINPVNRIIDITERSVLRGLDYDAKKEKLEFVNALEOP
 DVNPEPERAQAIGKERSLEAVVDLPDMKGLGVEDVGLLEAVEADEPTEVNAVE
 LDALAEKLAQLEFRFMSGEYSDADYLDIAGSGGNEADMSLEKRLMAESG
 FKTEIESEGEVAGIKSVTIKISGDVAGVGLPREYVHLVRSKSPEDSGRRHSES
 SAFVPEYDDIDIDEINPADLRIDVYRASGGQHVNRSAVATHIDPIGIVTOCON
 DRSHKNDKMDKMKAKIVTELEKQNAERQAEQKSDIGGSOITRSVYLDSEIRIK
 DLRGVEYTRNGAVLDSLDQETIASLAKGL"
 complement(6022..7755)
 /gene="recJ"
 /locus_tag="c3472"
 /complement(6022..7755)
 /gene="recJ"
 /locus_tag="c3472"
 /EC_number="3.1.-.-"
 /function="enzyme; Degradation of DNA"
 /note="Escherichia coli K-12 ortholog: b2892; Escherichia coli O157:H7 ortholog: z4230"
 /codon_start=1
 /transl_table=11
 /product="single-stranded-DNA-specific exonuclease recJ"
 /protein_id="AAN81920.1"
 /db_xref="GI:26109715"
 /translation="MKQIQLRRREYDETADLPALPPLRLRLVARSQVSOEERS
 VKMDLPQQLSGEKAVEITLVNAFRGRTIIVVGFEDADGATSTALVYAMSIGCSN
 IDLVNRRFEDAGLSEVVDQAHAGAOITVYNGISSHAGVHARSIGIPVYTD
 HILPBDTLPAAEAIINPNDNCFSPKSLAGVAFYMLAIFRTELPGQWFDGRTA
 IMLAEILDVAGVADVVPDANNRILITGQMSRIRAKRCPKIKALEVANDPO
 KLAASDGLPALGRLNAGRDMDGVCVALLCDNIGEARVYANLDAINTREIEIO
 GHOVEALITCEKLEERSRDLTPGCIAMTHHEMAGVGVILASIKERFRPVYAPAG
 DGTILGSGSLIGLHNRDLERLDLTVPEMLKFGHMAAGLSLEEDKFEIPOORFG
 ELYTERLDLSLQGEVSDGPIISAEMTMEVAQLRDAGPMGMPPEPFLFGHPRILLO
 QILVGRHLKVAVEVEVGGPPLDGIAFVVDALMPDNGVREVQVLAUKIDINEFGNS
 IQIILINDIWPV"
 complement(7761..8471)
 /gene="asbc"
 /locus_tag="c3473"

gene

CDS

/codon_start=1
 /transl_table=11
 /product="Isopentenyl-diphosphate delta-isomerase"
 /protein_id="AAN81915.1"
 /db_xref="GI:26109710"
 /translation="MTEHVLILNAQGVPTGLEKVAHATDTLLHLAFSSWLENKQ
 QLVTRALSKKAMPVGTNSVCGHPIQIGSENEAVIRRCRYELGYETPESITYDF
 MVMVQATNREARKLSAFLOUK"
 complement(2550..3260)
 /insertion_seq="IS1541A-like"
 /complement(2657..3166)
 /locus_tag="c3468"
 /complement(2657..3166)
 /locus_tag="c3468"
 /note="Residues 18 to 169 are 94.07 pct identical to residues 1 to 152 of 152 from SwissProt.40 : >sp1057334|T200_SALTY Transposase for insertion sequence element IS200"
 /codon_start=1
 /transl_table=11
 /product="Hypothetical protein"
 /protein_id="AAN81916.1"
 /db_xref="GI:26109711"
 /translation="MMSGKCCOTRQKQVPMGNKSLAHTFMNCKYHIVFAFKYRQ
 VYRERARRAIGCILKLCBWKSVRLIECCADHIMLVLEIPKMSVSGFMGLKGS
 SIMPYQEGDDLFKTRNREWCGRYVDTVGKNTAKIDYIKHOLEEDKMGHQLSPY
 RPSFTGRK"
 complement(3306..4823)
 /gene="lysS"
 /locus_tag="c3469"
 /complement(3306..4823)
 /gene="lysS"
 /locus_tag="c3469"
 /EC_number="6.1.1.6"
 /function="enzyme; Aminoacyl tRNA synthetases, tRNA modification"
 /note="Escherichia coli K-12 ortholog: b2890; Escherichia coli O157:H7 ortholog: z4228"
 /codon_start=1
 /transl_table=11
 /product="lysY tRNA synthetase"
 /protein_id="AAN81917.1"
 /db_xref="GI:26109712"
 /translation="MSEQHAQADAVVDLNNELKTRREKIANLREGIAPNPNRRDH
 TSDIHAFFGKNEELALNIEVAVAGRMTRIRKSKASFTYLDGGRIQLVAD
 DLPEGVNEQPKKMDLIGAKGKLEKTKTGETSTCTELRLTALRPLPKFHL
 OODAPRYRORYLIDISNDSRNTKVASQLISGRPMVNRGMEVETPMQVYIPGGL
 AARFETHHNAALIDMYLRLAPELYLKRLVGVGEYFELINRNENGLISVRNPEET
 MMELVMAVYKDLILELTSDFRTLADILKGTVEYGVDTLDGKPFKLTREALK
 KYREIDMADLDNDSAKAIAESIGIHVESWGLGIVTEIFEVAAHLIOTPTTTE
 YPAVESPLARRNVNPEITDRFEFFIGGRIGSGFSLINDAEOAORFLDOVAKAG
 DDEAMFYDEDYVALAHEGLPPTAGLIGIDRMVLMETNSTITIDVILFPAMREVK"
 3423..4838
 /locus_tag="c3470"
 3423..4838
 /locus_tag="c3470"
 /codon_start=1
 /transl_table=11
 /product="Hypothetical protein"
 /protein_id="AAN81918.1"
 /db_xref="GI:26109713"
 /translation="MFOCGDITFVVEHRRVYVICRGNLIQDALRIIFRVIOFAKAVI
 DPTNNEELKAVCDPRVNYTTRORNRFCITSNRGLDMQCFHFEDLQNDIADPT
 APLVNDADPFSNRRFREVQJRHVSFERHFLDRFTGAFERFEVQRYVAVGHRL
 TENILQSTQERFGOLDLFLVCISVSHFERHGLWMVYGNFVEVAVDEPYLTAT
 DHTLEVOQRDQIVAVEIOIYVMSDKRSTGARDHHRFNHKAAYHHELANTR
 EDLRHFKGVAFIVADEIOLITLTAFLILQAVETIORQCFQOQTOIGVDSQFA
 HGGEFOFPRADIAVEPEFELVYNARQIYTVGVQVQATNANVLDGNERRTTHRTG
 HHAASNGEDVGRFOFVILFTYFECVQVLRGMLATEIVREGNPLITVOBOLLTRFQF
 IYKIDYRSALCVLFRHVGSS"
 complement(4833..5931)
 /gene="prfb"

CDS

CDS

/locus_tag="c5622"
 /complement(join(4833..5855,5857..5931))
 /gene="prfb"
 /locus_tag="c5622"
 /function="factor; Proteins - translation and modification"
 /note="Rf-2: Escherichia coli K-12 ortholog: b2891; Escherichia coli O157:H7 ortholog: z4229; ribosomal slippage: an in-frame uga termination codon is located within the prf sequence, and a naturally occurring +1 frameshift is required for synthesis of Rf-2"
 /codon_start=1
 /transl_table=11
 /product="peptide chain release factor 2"
 /protein_id="AAN81919.1"
 /db_xref="GI:26109714"
 /translation="MFEINPVNRIIDITERSVLRGLDYDAKKEKLEFVNALEOP
 DVNPEPERAQAIGKERSLEAVVDLPDMKGLGVEDVGLLEAVEADEPTEVNAVE
 LDALAEKLAQLEFRFMSGEYSDADYLDIAGSGGNEADMSLEKRLMAESG
 FKTEIESEGEVAGIKSVTIKISGDVAGVGLPREYVHLVRSKSPEDSGRRHSES
 SAFVPEYDDIDIDEINPADLRIDVYRASGGQHVNRSAVATHIDPIGIVTOCON
 DRSHKNDKMDKMKAKIVTELEKQNAERQAEQKSDIGGSOITRSVYLDSEIRIK
 DLRGVEYTRNGAVLDSLDQETIASLAKGL"
 complement(6022..7755)
 /gene="recJ"
 /locus_tag="c3472"
 /complement(6022..7755)
 /gene="recJ"
 /locus_tag="c3472"
 /EC_number="3.1.-.-"
 /function="enzyme; Degradation of DNA"
 /note="Escherichia coli K-12 ortholog: b2892; Escherichia coli O157:H7 ortholog: z4230"
 /codon_start=1
 /transl_table=11
 /product="single-stranded-DNA-specific exonuclease recJ"
 /protein_id="AAN81920.1"
 /db_xref="GI:26109715"
 /translation="MKQIQLRRREYDETADLPALPPLRLRLVARSQVSOEERS
 VKMDLPQQLSGEKAVEITLVNAFRGRTIIVVGFEDADGATSTALVYAMSIGCSN
 IDLVNRRFEDAGLSEVVDQAHAGAOITVYNGISSHAGVHARSIGIPVYTD
 HILPBDTLPAAEAIINPNDNCFSPKSLAGVAFYMLAIFRTELPGQWFDGRTA
 IMLAEILDVAGVADVVPDANNRILITGQMSRIRAKRCPKIKALEVANDPO
 KLAASDGLPALGRLNAGRDMDGVCVALLCDNIGEARVYANLDAINTREIEIO
 GHOVEALITCEKLEERSRDLTPGCIAMTHHEMAGVGVILASIKERFRPVYAPAG
 DGTILGSGSLIGLHNRDLERLDLTVPEMLKFGHMAAGLSLEEDKFEIPOORFG
 ELYTERLDLSLQGEVSDGPIISAEMTMEVAQLRDAGPMGMPPEPFLFGHPRILLO
 QILVGRHLKVAVEVEVGGPPLDGIAFVVDALMPDNGVREVQVLAUKIDINEFGNS
 IQIILINDIWPV"
 complement(7761..8471)
 /gene="asbc"
 /locus_tag="c3473"

CDS

gene

CDS

Query Match 26.2%; Score 296.8; DB 1; Length 303121;
 Best Local Similarity 56.9%; Pred. No. 1.4e-73;
 Matches 566; Conservative 0; Mismatches 422; Indels 6; Gaps 1;
 119 TGGCTAAACCCCAAGTGAATCTTACGACACAGGAGGAGACCATCGCTGCGGGAAT 178
 69202 TGGCATTAACCCCAATTCACCAATTTTAGCAACCGCGGACCATTCGCGGTGCTACT 69143
 179 CTAGCGTAAAGNAGTACTCTGCTGAGAGCATCACCGTTGATMACTTCTTGACACCG 238
 69142 CCGCAACCAATCTTAACACTACACAGCGGTAAGTGGCGTAGAAATCTGTTAATGCG 69083
 239 TCCTGCGCATCAAGCACTTACCAACCAAGGCTGAACAGATCTCAAGCATTTGGCTCC 298
 69082 TGCCGCAACTGAAGACATTCGCAACTTAAAGCGCAGGAGTAGTGAATATTGGCTCC 69023
 299 AAGGATGACGGGTAGGTGTGCTTAACTAGCAAGCGCTGCATAGAGCTCTCTGCCCC 358
 69022 AGGACATGAACGATGATGTCTGCTGACTGCGCAAAAAAATTA-----CACCGACT 68969

gene

QY 359 AAAAGAGCCGAGCCGTGATCATCAACCGTGAAGTGAACACCATGAGAGACCGCTT 418
 Db 68968 GCGATTAACCTGAGCGCTTCTCTCTTACCCAGCTACCGACATGGAAGAAACCGCTT 68909
 QY 419 TCTTCTCAACCTGAGCGTGAAGAAACCGTCTCTCTGTAAGCGCCATGCGTC 478
 Db 68968 ACTTCTGAGCTGAGCGTGAAGAAACCGTCTCTCTGTAAGCGCCATGCGTC 68849
 QY 479 CAGGCTTCTCAATGAGTCTATGAGCGCCATGATCTCTTAACCGCGTGAATATGGA 538
 Db 68848 GGTTCACCTCTATGAGCGTGAAGAAACCGTCTCTCTGTAAGCGCCATGCGTC 68789
 QY 539 TCAACAAAGCCCTCTCTCAACAAAGAGTGTATGTAAGAGATGATGATGATGATG 598
 Db 68788 CTGATTAACCGCT 68729
 QY 599 CCAG 658
 Db 68728 GCGGATGATGAG 68669
 QY 659 GTAAATGCGCAGCTCTATGATGATGATGATGATGATGATGATGATGATGATG 718
 Db 68668 GTCTCTGAGTGTATGATGATGATGATGATGATGATGATGATGATGATGATG 68609
 QY 719 ACACCTTGTCAAGTGTATGATGATGATGATGATGATGATGATGATGATGATG 778
 Db 68608 ACACCGAG 68549
 QY 779 TTTTACGCTGAGCGCTGATGATGATGATGATGATGATGATGATGATGATGAT 838
 Db 68548 TTTTAACTGATGATGATGATGATGATGATGATGATGATGATGATGATG 68489
 QY 839 AAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 898
 Db 68488 ATGCGATCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 68429
 QY 899 TTTTAACTGATGATGATGATGATGATGATGATGATGATGATGATGATG 958
 Db 68428 TTTTAACTGATGATGATGATGATGATGATGATGATGATGATGATGATG 68369
 QY 959 CACCGAG 1018
 Db 68368 CTACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 68309
 QY 1019 TCAACCTGAG 1078
 Db 68308 TGAACCGAG 68249
 QY 1079 AGGCGATGAG 1112
 Db 68248 AGCAGATGAG 68215

RESULT 12
 LOCUS U32758 10173 bp DNA linear BCT 29-MAY-1998
 DEFINITION Haemophilus influenzae Rd section 73 of 163 of the complete genome.
 ACCESSION U32758 L42023
 VERSION U32758.1 GI:1573747
 KEYWORDS
 SOURCE Haemophilus influenzae Rd
 ORGANISM Haemophilus influenzae Rd
 Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
 Pasteurellaceae; Haemophilus.
 1 (bases 1 to 10173)
 Fleischmann, R.D., Adams, M.D., White, O., Clayton, R.A.,
 Kirkness, E.F., Kerlavage, A.R., Bult, C.J., Tomb, J., Dougherty, B.A.,
 Merrick, S.M., McKenney, K., Sutton, G.G., Fitzhugh, W., Fields, C.A.,
 Gocayne, J.D., Scott, J.D., Shirley, R., Liu, L.I., Glodek, A.,
 Kelley, J.M., Weidman, J.F., Phillips, C.A., Spriggs, T., Hedblom, E.,
 Cotton, M.D., Utterback, T., Hanna, M.C., Nguyen, D.T., Saudek, D.M.,
 Brandon, R.C., Fierman, J.L., Fritchman, J.L., Fritchman, J.L.,
 Geoghegan, N.S., Gnehm, C.L., McDonald, L.A., Small, K.V., Fraser, C.M.,

TITLE Smith, H.O. and Venter, J.C.
 JOURNAL Whole-genome random sequencing and assembly of Haemophilus
 MEDLINE Influenzae Rd
 PUBMED Science 269 (5223), 496-512 (1995)
 95350630
 7542800
 2 (bases 1 to 10173)
 Tatusov, R.L., Mushegian, A.R., Bork, P., Brown, N.P., Hayes, W.S.,
 Borodovsky, M., Rudd, K.E. and Koonin, E.V.
 Metabolism and evolution of Haemophilus influenzae deduced from a
 whole-genome comparison with Escherichia coli
 CURR. BIOL. 6 (3), 279-291 (1996)

TITLE
 JOURNAL
 MEDLINE
 PUBMED
 96398784
 8805245
 3 (bases 1 to 10173)
 White, O., Clayton, R.A., Kerlavage, A.R. and Fleischmann, R.D.
 Direct Submission
 Submitted (25-JUL-1995) The Institute for Genomic Research, 9712
 Medical Center Dr, Rockville, MD 20850, USA
 4 (bases 1 to 10173)
 White, O., Clayton, R.A., Kerlavage, A.R. and Fleischmann, R.D.
 Direct Submission
 Submitted (27-SEP-1997) The Institute for Genomic Research, 9712
 Medical Center Dr, Rockville, MD 20850, USA
 The H. influenzae sequence has been updated by R. Fleischmann. New
 database matches have been assigned, product names have been
 improved, and a number of frame shifts have been corrected. We
 gratefully acknowledge the work of Tatusov et. al. We have
 incorporated their annotation into the /notes fields of the
 corresponding H. influenzae genes
 5 (bases 1 to 10173)
 White, O., Clayton, R.A., Kerlavage, A.R., Fleischmann, R.D.,
 Peterson, J., Hickey, E., Dodson, R. and Gwin, M.
 Direct Submission
 Submitted (28-MAY-1998) The Institute for Genomic Research, 9712
 Medical Center Dr, Rockville, MD 20850, USA
 The whole genome was shifted by 588 nucleotides for a new start
 On Sep 30, 1996 this sequence version replaced gi:1221457.

REMARK
 COMMENT
 FEATURES
 source

gene
 CDS
 gene
 CDS
 gene
 CDS

/organism="Haemophilus influenzae Rd"
 /mol_type="genomic DNA"
 /db_xref="taxon:71421"
 247..1704
 /gene="HI0740"
 247..1704
 /note="HI0740"
 /note="similar to GB:M34393 SP:P18159 PID:142994
 PID:2226139 GB:AL009126 percent identity: 38.72;
 identified by sequence similarity: putative"
 /codon_start=1
 /transl_table=11
 /product="phosphomannomutase (yhx3)"
 /protein_id="AAC22400.1"
 /db_xref="GI:1573748"
 /translation="MGMRVLYVQAGAGLAELKGYDEPSIVIGYDGRKNSDYFARD
 TAEIMAGVQAYALIDPKLPIPVQALAYOYDTAGVAVTASHNPEDNGKYSVAGRA
 NCGGQIVSPADKTDALIDKYAAGRIODLPDSNDVYVNDVAVATITKTSIAKEPA
 CDINVTYAMGSGVTEVLSKLARKAGLQPHYADQVWPDCSTPPVNPNEEGALD
 LAIRVAKERNAEFLIANDPDADRLAVAVDQGMKSLHGVVGCFLGWTYAKOYQK
 OGTLACSLVSPALAEIARKYSFQSEETLTGFKYIGKVSGLFGEELGLVDPDKY
 RDKDGSAAIVFLDVRNLKKGKTLADYADEFTKEFGAVYSGQISIRVSDLSIEGKL
 MVALRNPPAEIAGYVNOEFDHITKIDROSIDLYVNLNNGRLIARPSGTBPKIKFYL
 DARGKPDADRYLAEPDEGVNHIIRQADYKQDC"
 1853..2182
 /gene="HI0742"
 1853..2182
 /note="HI0742"
 /note="hypothetical protein; identified by Genemark;
 putative"
 /codon_start=1
 /transl_table=11
 /product="H. influenzae predicted coding region HI0742"

/protein_id="AAC22407.1"
 /db_xref="GI:1573749"
 /translation="MGNLVAECRCICELNNGWTENTYTSVPPTLANNAPYGSVEN
 GKISSAIPAEIAEAADYAAVALVOYGEFGAGLTAQMDMTSGALTCCKRWYLY
 IKNHLHS"
 complement(2249, .2758)
 /gene="H10743"
 /note="similar to SP:P15040 GB:M24489 PID:147797
 PID:466747 GB:U00096 percent identity: 56.16; identified
 by sequence similarity; putative"
 /codon_start=1
 /transl_table=11
 /product="protein-export protein (secb)"
 /protein_id="AAC22401.1"
 /db_xref="GI:1573750"
 /translation="MSEOKDVAATEEQPVLOIQRIVYKDVSEAFENLPHIFOQEWK
 PKIGFDISTEYGDVLYVNIIVETLLESDGDAFICIEVROAGVFITSGEDVO
 MAHLISQCPNMLFYPARELYSNLVNNGTEPALMLSPVNEDELFVEYMNROAENAE
 KSEEQYRK"
 complement(2772, .3218)
 /gene="H10744"
 /note="similar to SP:P37688 PID:466749 GB:U00096
 PID:179040 percent identity: 50.74; identified by
 sequence similarity; putative"
 /codon_start=1
 /transl_table=11
 /product="conserved hypothetical protein"
 /protein_id="AAC22402.1"
 /db_xref="GI:1573751"
 /translation="KQEFISMATEFAOKHTLAVSWFAIFVNIYTFYKATSKFKYI
 THNEVIRLINSDEAIVDLRSLERFQGHILINSIVLPSKIKONICKESHKENALI
 LVDNCTGSASASVFLTKQFNSVFLKEGLSAMVAANLPLVKKK"
 /gene="H10745"
 /note="similar to GB:M34234 SP:P08085 GB:M34277 PID:145277
 PID:146597 percent identity: 70.52; identified by sequence
 similarity; putative"
 /codon_start=1
 /transl_table=11
 /product="L-asparaginase II (ansb)"
 /protein_id="AAC22403.1"
 /db_xref="GI:1573752"
 /translation="MKITKLALCTLFGLGVSIAMADLPNITLITATGTIAGSGSSY
 NSAYRAGSLIDTLEAVPEMKRNIANKGEOLIVKISOMNDEVMILKAIINAOCKS
 TGFVTHGTDMETAYFLDLTKCEKFPVVGAMRPATERSADGLILXYAVYAA
 DKSSGRGLVAMNNEVIGARDYTKSTTAQVDEHSPNGSGCYTINSVYDERSPES
 KRTIMPEVVERKIDSLPKYGIYASNAVEPLNALLNAGYGTISAGVGNVNAH
 LDRLEKAADSVYVYRSPVPTGYITTRDAEVDKSYFVASGLTNQKARVYLQALTL
 QTRKRVYIQQIEDF"
 /gene="H10746"
 /note="similar to GB:U14003 SP:P14409 GB:M27058 GB:X79886
 PID:146047 percent identity: 42.59; identified by sequence
 similarity; putative"
 /codon_start=1
 /transl_table=11
 /product="anaerobic C4-dicarboxylate membrane transporter
 protein (dcub)"
 /protein_id="AAC22404.1"
 /db_xref="GI:1573753"
 /translation="WSAMFLLQFAIVLCIFGARVNGIGLVEFGGLIALISFEGFL
 KPAQPIDVEMIVAVAAAMQAAGLIDYMIKATILIRNPHITFIPIAAVAVWLF
 TLAGTGIVAVSVLIVAEVRONGIRPERPSMAVIVASQFAIVASPIIAAVAVVAF
 LEPQIHLDVILMTVIPSITILGLFLACLFVNKMKELDDPEYORLMDPKRADIFAS
 TTSYKEVSVSTAKISVSLFLGALLVYLKGAASIRIRVFDKPMGMAHTTITMLSI

GALLITCKPDGTATTKGSVPHAGNRAVIAIFGLAMLDGTLMAHITTEKVKGLVE
 TAPWAFALFVLVSVNSOGATVATLELTKGATIGTIPAPVIGVAVNAGFFELPNC
 PLIASIDPDTGTETKIGKYLINHSFMLPGLISMFFCLLIGLSHVIL"
 6175, .7509
 /gene="H10747"
 /note="similar to SP:P00393 PID:581140 GB:U00096
 PID:1651546 PID:1787352 percent identity: 57.77;
 identified by sequence similarity; putative"
 /codon_start=1
 /transl_table=11
 /product="NADH dehydrogenase (ndh)"
 /protein_id="AAC22405.1"
 /db_xref="GI:1573754"
 /translation="MKNVIVYGGAGGAGGIELATFLGNKLRGROKAKVTLVDNATHMK
 PLHEIATGVMDVDGVSLSYRAHGNHFESEOSIIRINEQKYVELLAPYQGEEM
 LVIRARIPYDLYVIAIGSKSNDEMTKCAVNCIFLDSKALRPOHKLLELFESN
 RAUIDIGEEPRKQIVDENKYNVIAIVGGATGVLELHHAEDLSYGGKIDSSC
 LOYTLVAGRTFLPALDENLSAAVLDELEKENGIVOLNMTTEAOPNTLYIKDGEIK
 ADLIVMAVRASTVTOFQGLEINRINQVLKDTQTYVDSIFAIGDCAALIOSNG
 KLVPRQAQAHQMAKACAKNIFLFEKPKPLSKFYNDKGLVLSNFTALGSLNFKG
 KNPITVOGKFAQFAVYSLYRHHQHALGCIKIGLILVDKLNRYLKRLKH"
 complement(7589, .10021)
 /gene="H10748"
 /note="similar to SP:P00482 GB:K00127 PID:147295
 PID:148102 PID:148109 percent identity: 57.21; identified
 by sequence similarity; putative"
 /codon_start=1
 /transl_table=11
 /product="glycerol-3-phosphate acyltransferase (plsB)"
 /protein_id="AAC22406.1"
 /db_xref="GI:1573755"
 /translation="MAFIMYRQLSLPLSALVKNPPIPNPIEELSLNIHQPIYV

Query Match 26.28; Score 296.4; DB 1; Length 10173;
 Best Local Similarity 57.18; Pred. No. 1; e-73;
 Matches 562; Conservative 0; Mismatches 417; Indels 6; Gaps 1;

QY 126 ACCCCAAGTACTATCTTACCCAGAGGACCCAGTCTGGGGGAATCTAGCT 185
 DB 3515 ACCCAATATTACATCTTGGCAACGGGTGTACCATTCAGAGCGCAAGTTCGT 3574
 QY 186 CAAGAGTACTCTCTGCTGAGCAGACGACCTGATTAACCTCTTGCACCCCTGCTC 245
 DB 3575 AATTTCTGCTATTAAGCTGTGACACATTAATGATTAATCTTAAATGAAGCTGTACACAGA 3634
 QY 246 CATCAACGACCTTACCCAGGAGTCAAGAGTCTCAAGTCTGCTCCCAAGAT 305
 DB 3635 AATGAAATATTGTCACCACTTAAAGGTGAGCAATTTGTAANAATGTTCAACAAGACAT 3694
 QY 306 GACGGTAAAGTGTGGCTTAACTAGCCAAAGCTGTCATGAGCTCTGCGCCCAAAAAGA 365
 DB 3695 GATGACGAAGTGTGCTTAAAGTGTGACCAAGCCATCAATATGCTC-----AATGTAAAG 3748
 QY 366 GACCGAAGCTGTATCATCCCATGAACTGACACATGAGAGAGACCGCTTCTTCT 425
 DB 3749 TACTGATGATTTGCTATACCATGTGTACAGTATACGAGAAAGACGCGTTATTCTT 3808
 QY 426 CAACCGACGGTGAAGCAAAAACCTGTGCTTGTAGGCGCATCGCTCCAGGCTC 485
 DB 3809 AGATTTCACGTAATATGAAAAACGGTGTCTTCGTTGGGGCAATCCGCTGCAAC 3868
 QY 486 TTTCATGATGCTGTATGCGCCCATGATCTATTAAGCGCGTGAATGTGAGATCAACA 545
 DB 3869 AAAAAAAGTGTGATGCGCCATTAATCTTACAAATGCTGCTGTGCGACGACAGACA 3928
 QY 546 ACCCTCTACTAACAAAGAGTGTGATGTGATGAAGATGATTCACGCGCCAGAGA 605
 DB 3929 AAAATCAAGTGTGCTGTGTTTGTGCGCATATATGAAGTACTAGTGTGCGGA 3988


```

QY 606 AGCGACCAAGCTCAACACCCGAGTCAATGCTTCTGCCCCACACAGTAAT 665
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3989 TGTACAAAACACATAGACGAGTCAACGTTCCATTCCAAATTAATGTTCT 4048
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 666 CGGACAGCTCTATTGGAAGTGAATTCATCTCACTCAATCCCTTCACCTCACACCT 725
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4049 AGGCTATATTCATTAACAGCAAGTGAATTAACGTTCCACCAAGCAACATATCAT 4108
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 726 TCGAAGTAGTGTGATTAATGCAAAATGCAAGAACTCCCGAGATGATTTCTTACGC 785
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4109 CAACACTCGTTTAAGTGAAGAAATTAATAGCTACCAAGAGTGGGATTTATTAATGC 4168
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 786 TCACCCCGATGATGATGATTTAGTCAATGCAAGCCCTTCAGGAGGCAAGGAAT 845
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4169 TTAATCAAAATGCACTGTGAAACATTAACGATTAATCAATGCTGATCAAGGAT 4228
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 846 CATTCATGAGGATGAGCAATGGAACCTTTCCTTTCAGTCAAAATGCTCTGANA 905
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4229 TGTATCTGAGAGTGTGCAATGGAATGTAATGCTGACACTTATGATCGCTTGAANA 4288
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 906 AGCAGCCAAATCAAGCGTATGCTGCTGGAAGCTTAAGTGGGAGTGTCCACAC 965
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4289 AGCGGCAAAAGATAGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4348
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 966 CCAAGAGGCTGAGTGTGATGATTAAGAACTGTTTGTGCTACAGAGTCTCAACCC 1025
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4349 TGTGAGAGCTGAGTGTGATGATTAAGAACTGTTTGTGCTACAGAGTCTCAACCC 4408
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1026 TCAAAAAGCAGAGTGTGATGATTAAGAACTGTTTGTGCTACAGAGTCTCAACCC 1085
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4409 ACAGAAAGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4468
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1086 CCAAAAGATCTCTCCACCTATTA 1110
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4469 TCAACATATTTTCAAGACTTTCA 4493
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 13
AR274513_07

WPCOMMENT
Sequence split into 19 fragments LOCUS AR274513 Accession AR274513

Fragment Name	Begin	End
AR274513_00	1	110000
AR274513_01	100001	210000
AR274513_02	200001	310000
AR274513_03	300001	410000
AR274513_04	400001	510000
AR274513_05	500001	610000
AR274513_06	600001	710000
AR274513_07	700001	810000
AR274513_08	800001	910000
AR274513_09	900001	1010000
AR274513_10	1000001	1100000
AR274513_11	1100001	1210000
AR274513_12	1200001	1310000
AR274513_13	1300001	1410000
AR274513_14	1400001	1510000
AR274513_15	1500001	1610000
AR274513_16	1600001	1710000
AR274513_17	1700001	1810000
AR274513_18	1800001	1910000

Continuation (8 of 19) of AR274513 from base 700001 (AR274513 Sequence 1 from patent US

Query Match 26.2% Score 296.4; DB 6; Length 110000;
Best Local Similarity 57.1%; Pred. No. 1.5e-73;
Matches 562; Conservative 0; Mismatches 411; Indels 6; Gaps 1;

```

QY 126 ACCCAAGTACTATCTAGACCAAGAGGACCAATCGCTGTGGGGGAATCTAGGT 185
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 102722 ACCCAATATTAATCTAGACCAAGAGGACCAATCGCTGTGGGGGAATCTAGGT 102781
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 186 CAAGACTACTACTCTGCTGGAGGAGTCACTGATAGAGTCTTTCAGAGCCGCTCC 245
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

Db 102782 AATTTCTGCTATTAAGCTGGAATTAATTAATTAATTAATTAATTAATTAAT 102841
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 246 CATCAAGAGCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 305
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 102842 AATTAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 102901
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 306 GACGGGTAAAGTGTGCTTAACTAGCAAGCGTGTCAATGAGTCTGCTGCTGCTGCT 365
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 102902 GAATGAGCAAGTGTGCTTAACTAGCAAGCGTGTCAATGAGTCTGCTGCTGCT 102955
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 366 GACCGAAGCGTGTATCAATCAATCAATCAATCAATCAATCAATCAATCAATCA 425
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 102956 TACTGATGATTTGCTATTAATCAATCAATCAATCAATCAATCAATCAATCA 103015
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 426 CAACCTCAGCTGTAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 485
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 103016 AGATTTAACTGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 103075
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 486 TTCCATGAGTGTGATGAGTGTGATGAGTGTGATGAGTGTGATGAGTGTGATGAG 545
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 103076 AGAAAAAGTGTGATGAGTGTGATGAGTGTGATGAGTGTGATGAGTGTGATGAG 103135
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 546 AGCTCTCTCTAACAAGAGTGTGATGAGTGTGATGAGTGTGATGAGTGTGATGAG 605
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 103136 AATTAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 103195
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 606 AGCGACCAAGCTCAACACCCGAGTCAATGCTTCTGCTGCTGCTGCTGCTGCTGCT 665
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 103196 TGTACAAAACACATAGACGAGTCAACGTTCCATTCCAAATTAATGTTCTCT 103255
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 666 CGGACAGCTCTATTGGAAGTGAATTCATCTCACTCAATCCCTTCACCTCACACCT 725
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 103256 AGGCTATATTCATTAACAGCAAGTGAATTAATGATGATGATGATGATGATGAT 103315
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 726 TCGAAGTAGTGTGATTAATGCAAAATGCAAGAACTCCCGAGATGATTTCTTACGC 785
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 103316 CAACACTCGTTTAAGTGAAGAAATTAATAGCTTACCAAGAGTGGGATTTATTAATGC 103375
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 786 TCACCCCGATGATGATGATTTAGTCAATGCAAGCCCTTCAGGAGGCAAGGAAT 845
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 103376 TTAATCAAAATGCACTGTGAAACATTAACGATTAATCAATGCTGCTGCTGCTGCT 103435
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 846 CATTCATGAGGATGAGCAATGGAACCTTTCCTTTCAGTCAAAATGCTCTGANA 905
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 103436 TGTATCTGAGAGTGTGCAATGGAATGTAATGCTGCACTTAATGATCGCTTGAANA 103495
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 906 AGCAGCCAAATCAAGCGTATGCTGCTGGAAGCTTATAGTGGGAGTGTCCACAC 965
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 103496 AGCGGCAAAAGATAGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 103555
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 966 CCAAGAGCTGAGTGTGATTAAGAACTGTTTGTGCTACAGAGTGTCTCAACCC 1025
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 103556 TCGTGAAGCTGAGTGTGATTAAGAACTGTTTGTGCTACAGAGTGTCTCAACCC 103615
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1026 TCAAAAAGCAGAGTGTGATGATTAAGAACTGTTTGTGCTACAGAGTGTCTCAACCC 1085
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 103616 ACAGAAAGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 103675
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1086 CCAAAAGATCTCTCCACCTATTA 1110
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 103676 TCAACATATTTTCAAGACTTTCA 103700
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 14
AR274513_08

WPCOMMENT
Sequence split into 19 fragments LOCUS AR274513 Accession AR274513

Fragment Name	Begin	End
AR274513_00	1	110000
AR274513_01	100001	210000
AR274513_02	200001	310000
AR274513_03	300001	410000
AR274513_04	400001	510000

AR274513_05 500001 610000
 AR274513_06 600001 710000
 AR274513_07 700001 810000
 AR274513_08 800001 910000
 AR274513_09 900001 1010000
 AR274513_10 100001 1110000
 AR274513_11 110001 1210000
 AR274513_12 120001 1310000
 AR274513_13 130001 1410000
 AR274513_14 140001 1510000
 AR274513_15 150001 1610000
 AR274513_16 160001 1710000
 AR274513_17 170001 1810000
 AR274513_18 180001 1830121

Continuation (9 of 19) of AR274513 from base 800001 (AR274513 Sequence 1 from patent US

Query Match 26.2%; Score 296.4; DB 6; Length 110000;
 Best Local Similarity 57.1%; Pred. No. 1.5e-73;
 Matches 562; Conservative 0; Mismatches 417; Indels 6; Gaps 1;

QY 126 ACCCCAGTACTTCTAGCCACAGAGGACCACTCCCTGTTCCGGGGAATCTAGCGT 185
 DB 2722 ACCAAATATTACATCTTGGCAGAGGGGTGATACCATTCAGAGGCGGCAAGTCCGT 2781
 QY 186 CAAGAGTAGCTACTCTGCTGGAGCAGTCCGTTGATAGCTTCTTGACGCGTCCCTGC 245
 DB 2782 AAATTCGCGTAATAAGCTGACACATTAATGATGATCTTAATGAAAGCTGTACCGCA 2841
 QY 246 CATACAGACCTCAACCATCAAGGTTGAACAGATCTCAAGCATTTGGCTCCCAAGAGAT 305
 DB 2842 AATGAAAAATTTGCGCAACATTAAGGTGAGCAAAATGTAATAATGATGATCAAGCAT 2901
 QY 306 GACGGTAAAGTGGTGAATCAAGCAAGCGTCAATGAGTCTCTCGCCCAAAAAGA 365
 DB 2902 GAATGAGGAAGTCTGGCTAAACCTGGCAAAAGCCATCAATGCTC-----AATGTAAGA 2955
 QY 366 GACCGAAGCCGTGATCAACCATCAAGGTTGAACAGATCTCAAGCATTTGGCTCCCAAGAGAT 425
 DB 2956 TACTGATGAGTTTGTCTATACCATTCAGATGATGATGATGATGATGATGATGATGATGAT 3015
 QY 426 CAACCTCAAGGTTGAAAAAGCAAAACCTGCTGCTGTTGAGGCGCCATGCTCCAGGCTC 485
 DB 3016 AGATTTAACCGTAAATGATGAAAAACCGTTGCTGCTGTTGAGGCGCAATGCTCCAGGCTC 3075
 QY 486 TTCCATGAGTCTGATGAGCCCATGATCTCTTAACGCGGTAATGATGATGATGATGATGAT 545
 DB 3076 AGAAAAAGTGTGATGAGCCCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3135
 QY 546 AGCTCTACTAACAAGAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 605
 DB 3136 AAATTAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3195
 QY 606 AGGACCAAGCTCAACACACGCACTCAATGATCTCTTAACGCGGTAATGATGATGATGATGAT 665
 DB 3196 TGTAAACAAAACAGTAGACGCGCATGCAAGGTTCCATTCACCAAAATTAATGATGATGAT 3255
 QY 666 CGGCAAGCTCTATGAGCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 725
 DB 3256 AGGCTATATTCATTAACAGCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 3315
 QY 726 TGCAGTAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 785
 DB 3316 CAACACTCGCTTAACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3375
 QY 786 TCACCCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 845
 DB 3376 TTATTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3435
 QY 846 CATCATGAGGCAAGGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 905
 DB 3436 TGTATCTCGAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3495
 QY 906 AGCAGCCAAATCAAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 965

DB 3496 ACCCGCAAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3555
 QY 966 CCAAGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1025
 DB 3556 TGTGACGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3615
 QY 1026 TCAAAAAGCAGAGTCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1085
 DB 3616 ACAGAAAGCCGCGGCTCTTCAATTTAGCTTTAATCACTTAAGATCAAAAGTAT 3675
 QY 1086 CCAAAAGATCTCTCCACTATTA 1110
 DB 3676 TCACCAATATTTCGAAGACTTCTAA 3700

RESULT 15
 ECOLASNII 1530 bp DNA linear BCT 26-APR-1993
 LOCUS E.coli L-asparaginase II (ansb) gene, complete cds.
 DEFINITION M34277.1 GI:146596
 VERSION M34277.1 GI:146596
 KEYWORDS L-asparaginase II.
 SOURCE Escherichia coli
 ORGANISM Escherichia coli
 Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Escherichia.
 1 (bases 1 to 1530)
 JENNINGS, M.P. and Beacham, I.R.
 Analysis of the Escherichia coli gene encoding L-asparaginase II, ansb, and its regulation by cyclic AMP receptor and FNR proteins
 J. Bacteriol. 172 (3), 1491-1498 (1990)
 PUBMED 90170867
 COMMENT 2407723
 FEATURES
 source
 Original source text: E. coli (strain K-12) DNA.
 location/Qualifiers
 1..1530
 /organism="Escherichia coli"
 /mol_type="genomic DNA"
 /db_xref="taxon:562"
 323..>1431
 /product="ansb mRNA"
 /evidence="experimental"
 385..1431
 /EC_number="3.5.1.1"
 /note="L-asparaginase II (ansb) precursor"
 /codon_start=1
 /transl_table=11
 /protein_id="AA24062.1"
 /db_xref="GI:146597"
 /translation="MEFRTALALVMEFSGAALAPNTTILATGTTAGGDSATK
 SNTVAVGEVENVNAVPOKDIANVGEVNVIGSDMNDVLTLLAKININDCKT
 DGFVTHGDVMEIETAVFLDLVYKCDKPVVAVGAMSPSMSDGPENLVNAVTPAD
 KASANGVAVVMDVLDGSDVYKNTTDDVAFKSVNGPLGYTHNGKIDYQVTPARK
 HSDPEVSKINELPKVGIYVAVNASDLPKAIYDADYDGVSGVNGNLTQSV
 DLTATATKGTAVVRSRVPFGATTTDADVDADKTFVASGTLNPKARVLQALTO
 TNDPQIDQIFNY"
 385..450
 /note="L-asparaginase II signal peptide"
 451..1428
 /product="L-asparaginase II"
 /EC_number="3.5.1.1"
 /base_count 409 a 381 c 382 g 358 t

sig_peptide
 mat_peptide
 BASE COUNT 409 a 381 c 382 g 358 t
 ORIGIN

Query Match 26.1%; Score 295.2; DB 1; Length 1530;
 Best Local Similarity 56.8%; Pred. No. 1.7e-73;
 Matches 565; Conservative 0; Mismatches 423; Indels 6; Gaps 1;

QY 119 TGGCTAAACCCCAAGTATCTTACCAAGAGGACCACTCGTGTGGGGGAAT 178
 DB 446 TGGCATTAACCAATATCACTTTAGCAACGCGGAGCCATTCGCGGTGTGTGACT 505

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd

OM nucleic - nucleic search, using sw model

Run on: September 4, 2003, 15:11:40 ; Search time 346 Seconds
(without alignments)
8839.489 Million cell updates/sec

Title:	US-09-937-982-3
Perfect score:	1133
Sequence:	1 atggcagcagcatcatca.....aagaaggygaatctcttcac 1133

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

```
Searched:      2552756 seqs, 1349719017 residues
Total number of hits satisfying chosen parameters: 5105512
```

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: /SDSL/gcgdata/genseq/genseqn-emb1/NA1980.DAT *

2: /SDSL/gcgdata/genseq/genseqn-emb1/NA1981.DAT *

3: /SDSL/gcgdata/genseq/genseqn-emb1/NA1982.DAT *

4: /SDSL/gcgdata/genseq/genseqn-emb1/NA1983.DAT *

5: /SDSL/gcgdata/genseq/genseqn-emb1/NA1984.DAT *

6: /SDSL/gcgdata/genseq/genseqn-emb1/NA1985.DAT *

7: /SDSL/gcgdata/genseq/genseqn-emb1/NA1986.DAT *

8: /SDSL/gcgdata/genseq/genseqn-emb1/NA1987.DAT *

9: /SDSL/gcgdata/genseq/genseqn-emb1/NA1988.DAT *

10: /SDSL/gcgdata/genseq/genseqn-emb1/NA1989.DAT *

11: /SDSL/gcgdata/genseq/genseqn-emb1/NA1990.DAT *

12: /SDSL/gcgdata/genseq/genseqn-emb1/NA1991.DAT *

13: /SDSL/gcgdata/genseq/genseqn-emb1/NA1992.DAT *

14: /SDSL/gcgdata/genseq/genseqn-emb1/NA1993.DAT *

15: /SDSL/gcgdata/genseq/genseqn-emb1/NA1994.DAT *

16: /SDSL/gcgdata/genseq/genseqn-emb1/NA1995.DAT *

17: /SDSL/gcgdata/genseq/genseqn-emb1/NA1996.DAT *

18: /SDSL/gcgdata/genseq/genseqn-emb1/NA1997.DAT *

19: /SDSL/gcgdata/genseq/genseqn-emb1/NA1998.DAT *

20: /SDSL/gcgdata/genseq/genseqn-emb1/NA1999.DAT *

21: /SDSL/gcgdata/genseq/genseqn-emb1/NA2000.DAT *

22: /SDSL/gcgdata/genseq/genseqn-emb1/NA2001A.DAT *

23: /SDSL/gcgdata/genseq/genseqn-emb1/NA2001B.DAT *

24: /SDSL/gcgdata/genseq/genseqn-emb1/NA2002.DAT *

25: /SDSL/gcgdata/genseq/genseqn-emb1/NA2003.DAT *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1129.8	99.7	1133	20	AAK03474	Molnella succinog
2	1129.8	99.7	1133	21	AAK62312	W. succinogenes as
3	300.4	26.5	1044	21	AAK82834	L-asparaginase DNA
4	296.4	26.2	1830	121	AAAT4063	Haemophilus influ
5	294	25.9	1848	17	AAAT6346	Chimeric gene conf
6	285.2	22.9	1196	19	AAK14328	H. pylori GHPD 187
7	223.2	19.7	2837	8	AAAT0557	Sequence encoding
8	215.8	19.0	1014	15	AAK06438	Pseudomonas glutam

9	199	17.6	1774	25	ABX268081
10	108.8	9.6	345	24	AAE60196
11	103.6	9.1	543	24	AA5211124
12	102	9.0	543	24	AA5211116
13	102	9.0	543	24	AA5211126
14	102	9.0	543	24	AA5211122
15	102	9.0	543	24	AA5211126
16	101.6	8.9	1436	25	ACCA4571
17	101.4	8.9	552	19	AAV13947
18	100.4	8.9	681	24	AA521115
19	100.4	8.9	681	24	AA521121
20	100.4	8.9	681	24	AA521125
21	100.4	8.9	682	24	AA521117
22	100.4	8.9	682	24	AA5211123
23	100.4	8.9	776	24	AA521119
24	98.4	8.7	1435	25	ACCA4570
25	96.4	8.5	543	24	AA521120
26	94.2	8.3	900	24	ABV73014
27	93.4	8.2	1247	20	AA529725
28	92.8	7.3	960	25	ABX07792
29	82.8	7.3	2162358	25	AB55645
30	79.8	7.0	316	25	AB580804
31	75.8	6.7	2365589	24	ABA90502
32	74	6.5	2231	19	AAE52398
33	72.6	6.4	366	22	AAE25008
34	71.4	6.3	969	22	AAH53259
35	71.4	6.3	981	24	ABN90934
36	71.4	6.3	4020	22	AAH54070
37	71	6.2	360	25	ABZ68085
38	70.4	6.3	1230	22	AAE25127
39	68.4	6.0	6682	20	AAH12993
40	68.4	6.0	6682	24	ABX98788
41	67.6	5.9	619	21	AAZ50042
42	67.4	5.9	654	21	AAZ50043
43	66.6	5.9	501	22	AAH06207
44	66.2	5.8	1455	22	AAE50134
45	66.2	5.8	2847	22	AAE25020

ALIGNMENTS

RESULT 1	
AA03474	
ID	AA03474 standard; DNA; 1133 BP.

AC AAX03474;

DT 30-APR-1999 (first entry)

DE Wolinella succinogenes L-asparaginase DNA

KW L-asparaginase; amplification; treatment; disease; asparagine depletion;
 KW malignant disease; haematology; lymphoma; leukaemia; myeloma; AIDS;
 KW autoimmune disease; rheumatoid arthritis; systemic lupus erythematosus;
 KW covalent modification; acylation; pharmacokinetics; immunogens; spleen;
 KW hypersensitive; reduced toxicity; immunosuppression; allergy; thymus;
 KW lymphocyte marker distribution; hepatotoxicity; ss.

05 *Wolinella succinogenes*

PN W09856410-A1.

PD 17-DEC-1998.

PF 09-JUN-1998;

PR 09-JUN-1997; 97US-0049085

PA (CHIL-) CHILDRENS HOSPITAL LOS ANGELES.

PI Durden DL;

Erwinia carotovora
Helicobacter pylori
 DNA encoding Tm 3
 DNA encoding Tm 2
 DNA encoding Tm 3
 DNA encoding Tm 3
 DNA encoding Tm 7
 pMV04836 glucose 1
Salmonella sefta ge
 DNA encoding Tm 2
 DNA encoding Tm 3
 DNA encoding Tm 7
 DNA encoding Tm 2
 DNA encoding Tm 7
 DNA encoding Tm 2
 DNA encoding Tm 3
 DNA encoding Tm 2
 DNA encoding Tm 1
 pMV04835 glucose 1
 DNA encoding Tm 1
 Recombinant SOAgp5
 Insert from pE28C
S. pneumoniae type
Streptococcus pneumoniae
Erwinia carotovora
 Genomic sequence c
Streptococcus pneumoniae
 Nucleotide sequence
S. epidemidis ope
Staphylococcus epidermidis gen
S. epidermidis gen
Erwinia chrysanthi
 Nucleotide sequence
Enterococcus faecalis
Enterococcus faecalis
 DNA encoding Hepat
 DNA encoding Hepat
Staphylococcus aureus
 Ltn-Bcl-Xl apopos
 Nucleotide sequence


```

Db      62 TGGCATTTACCAATTCACCAATTTTGAACACCGGGGACCAATGCGGTGGTACT 121
QY      179 CTAGCGTCAAGAGTACTCTGCTGGAGAGTACAGTGTATAGCTTCTTGACGCC 238
Db      122 CCGCACCAAACTCTACTACACAGCGGGTAAAGTTGGCGTGAATAATCTGGTTAAWGGG 181
QY      239 TCCCGTCATCAAGACCTAGCCACCATCAAGGGTGAACAGATCTCAAGCATTTGGCTCC 298
Db      182 TGGCGCACTAAAGACATTTGGAGACGTTAAAGCGAGACAGTATGATTAATCGGCTCC 241
QY      299 AAGAGATACGGGTAGGTGTGGCTTAACCTAGCCACAGCTGTCAATAGCTCTCGGCC 358
Db      242 AGGACATGAACGATATGTCTGGTACACCTGGCGAAAAAATTA-----CACCGACT 295
QY      359 AAAAGAGACCGAAGCCGTGATCATCACCCATGGAACCATGAGACCATGAGACCGCT 418
Db      296 GCGATMAAACCGAGCGCTTCGTCATTAACCGATGACGACACATGGAAGAAACCGCTT 355
QY      419 TCTTCTCACTCAAGCTGTAAGAAAGCCAAACCTGTCTGCTTGTAGCGCCATGCGTC 478
Db      356 ACTTCTCGACCTCAAGCTGTAAGAAAGCCAAACCGGTGATGGTGGCGCAATGGCC 415
QY      479 CAGCTCTTCCATAGTGTGATGGCCCATGAACTCTATTAAGCCCTGATGATGAGCA 538
Db      416 GGTCCAGCTCATAGCGAGACGCTCATTTCAACCTGATTAACCGGTGATGACCGGAG 475
QY      539 TCAACAAGCCTCTACTAACAAGAGAGTGTGATTTGATGATAGATGATGATGATGATG 598
Db      476 CTGATTAAGCGCTCGCTAATGCTGCGTGTGATGATGATGATGATGATGATGATGATG 535
QY      599 CCAAGAGAGGAGACCAAGCTCAACACCAACCGCATGATGATGATGATGATGATGATG 658
Db      536 GTCCGATGTCACCAACCAACCAACCGCATGATGATGATGATGATGATGATGATGATG 595
QY      659 GTAAATCGGACAGTCTATTATGCAAGAGTGTGATTTGATGATGATGATGATGATGATG 718
Db      596 GTCTCTGGGATATCTACACAGGTAAGATGATGATGATGATGATGATGATGATGATGATG 655
QY      719 ACACCTTGAAGAGTGTGATTTAGCAAAATGCAAGAACTGCCAGAGTGTGATGATG 778
Db      656 ACACCAAGAGTGTGATTTAGCAAAATGCAAGAACTGCCAGAGTGTGATGATGATG 715
QY      779 TTTACGCTCAACCGCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 838
Db      716 TTTATATCTACGCTAAGCAATCCGATCTTCGGCTTAAGCACTGATGATGATGATGATG 775
QY      839 AAGGAATCATCATGACGAGTGGCAATGGAAACCCCTTCCCTTGTACTCAAAATGCTC 898
Db      776 ATGGCATCGTTAGCGGTGTGATGATGATGATGATGATGATGATGATGATGATGATG 835
QY      899 TTTAAAAAGAGCAAAATCAAGCGGTAGTCTGCTGCAAGCTTGAAGTGGGCAAGTGT 958
Db      836 TGGCAACCGCGGAAAAAGCACTGCAATGAGCTTCTCCCGGTACCGACGCGGTG 895
QY      959 CCAACACCCCAAGAGGTGAAGTGAAGAACTTGGTTTGTGGTACAGAGATC 1018
Db      896 CTACCACTCAAGATGCTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 955
QY      1019 TCAACCTCAAAAGCCAGAGTGTGATTTAGTGAAGCTTCAACCAAACTAGTGAAG 1078
Db      956 TGAACCGCGGAAAAAGCGCGGTCTGCTGAGCTGAGCGCAAAACCAAGATCGCG 1015
QY      1079 AGCGATCCAAAGATCTTC 1098
Db      1016 AGCAGATCCAGCAGATCTTC 1035

```

RESULT 4
AAT42063
ID AAT42063 standard; DNA: 1830121 BP.
XX
AAT42063;

```

XX      14-SEP-1999 (first entry)
DT      Haemophilus influenzae complete genome sequence.
XX      DE
XX      KW Genome; bacterium; Haemophilus influenzae; computer readable medium;
XX      KW expression modulating fragment; regulation; gene expression; vector;
XX      KW organism; open reading frame; ORF; ds.
XX      OS Haemophilus influenzae.
XX      PN W09633276-A1.
XX      PD 24-OCT-1996.
XX      PF 22-APR-1996; 96WO-US05320.
XX      PR 07-JUN-1995; 95US-0487429.
XX      PR 21-APR-1995; 95US-0426787.
XX      PR 07-JUN-1995; 95US-0476102.
XX      PA (HUMA-) HUMAN GENOME SCI INC.
XX      PA (UYO ) UNIV JOHNS HOPKINS.
XX      PT Adams MD, Fleischmann RD, Smith HO, Venter JC, White O;
XX      PT WPI, 1996-485782/48.
XX      DR
XX      PT Haemophilus influenzae Rd genome recorded on computer readable
XX      PT medium - useful for identifying commercially important nucleic acid
XX      PT fragments by homology searching
XX      PS Claim 1; Page 77.2-77.1091; 1291pp; English.
XX      CC This sequence represents the complete genome sequence of the bacterium
XX      CC Haemophilus influenzae strain Rd. The invention relates to a computer
XX      CC readable medium (CRM) having recorded upon it the complete H. influenzae
XX      CC nucleotide sequence (1), a representative fragment of (1) or a nucleotide
XX      CC sequence at least 99% identical to (1). By providing the full-length
XX      CC genomic sequence in a computer readable form, it is possible to identify
XX      CC commercially important nucleic acid fragments and expression modulating
XX      CC fragments (EMFs) of the Haemophilus genome. The EMFs can be used to
XX      CC regulate the expression of a nucleic acid molecule. Vectors and altered
XX      CC organisms comprising the predicted ORFs can be used to produce any of the
XX      CC polypeptide fragments of the H. influenzae Rd genome.
XX      CC
XX      SQ Sequence 1830121 BP; 567399 A; 350615 C; 347389 G; 564036 T; 682 other;

```

Query Match 26.28; Score 296.4; DB 17; Length 1830121;
Best Local Similarity 57.18; Pred. No. 8.1e-81;
Matches 562; Conservative 0; Mismatches 417; Indels 6; Gaps 1;

```

QY      126 ACCCAAGTACTTCTCTAGCAAGAGGACCACTGCGGTGGGGAATGATACCT 185
Db      802722 ACCAATATATTAATCTTGGCAACGGGTGTGATGATGATGATGATGATGATGATG 185
QY      186 CAGAGTACTCTCTGAGACGATCACCGTTGATTAAGCTTTCAGACCGTCTGC 245
Db      802782 AATTTCTGCGTATTAACCTGACAAATTAATGATGATGATGATGATGATGATGATG 245
QY      246 CATCAAGACCTGCTGCAACCAAGGTAAGATGATCAAGATTTGGCTCCCAAGAT 305
Db      802842 AATGAAATATTTGCAACATTAAGGTGAGCAAAATTTGAAATTAAGTTCAACAAAGCAT 305
QY      306 GAGCGGTAGGTGTGCTTAAGTCAAGCGGTGATGATGATGATGATGATGATGATGATG 365
Db      802902 GAATGACGAGTGTGCTTAAGTCAAGCGGTGATGATGATGATGATGATGATGATGATG 365
QY      366 GACCGAAGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 425
Db      802956 TACTGATGATTTGTCTATTACCATGATGATGATGATGATGATGATGATGATGATG 425
QY      426 CAACCTCAGGTTAAAGCAAAACCAAAACCTGCTGCTTGTGAGGCGCATGCTCAGGCTC 485

```


Db	803016	AGATTTAAACCGTAAATATGTAAGAAAAACCGGTGTGTTCTCGTTGGGCAATATCGTCTGCAAC	803075
Oy	486	TTCCATGAGTCTGTATGGCCCCATGAATCTCTATTAACGCCGCTGAATGTATGGCATCAACA	545
Db	803076	AGAAAAAGTCTGTATGGCCCATTTAAATCTTACATGCTGTGCTGTGCGACAGACAA	803135
Oy	546	AGCCCTCTACTAACAAGAGATGTGATTTGTGATGAACGATAGATTCACGCGCCAGAGA	605
Db	803136	AAATATCAAGTGTGCTGTGCTGTTTATGTCGCAATGATTAATGAAGTACTAGTGCTGCGA	803195
Oy	606	AGCGACCAAGCTCAACACCAACCCGACGATCAATGATTTGCTTCCGCCACACAGATAAAT	665
Db	803196	TGTAAACAAAAACAGTATGACGCGCAGTGCAGAAAGCTTCCATCCACCAATTAATGTTCTCT	803255
Oy	666	CGGCATCTCTATTATGCGAAAGTCGAGATTTCACTCAATCCGTTGACCTCACACCT	725
Db	803256	AGGCTATTTCTCAATACAGCAAAAGTGACTATACAGTTCCCGAAGAAAGCAATATACAT	803315
Oy	726	TGCAAGTGAATTTGATTTATTAAGCAAAATCGAAGACTCCCGAGATCGATATTTCTTACGC	785
Db	803316	CAACACTCGTTTAACTGTAAGAAAAATAGATAGCCTACCCAAAGTGGGATTAATTTATGC	803375
Oy	786	TCACCCCGCATGATACTGATGTTTATGTCATATCAGCCCTTCACGCGAGGCCAAGAT	845
Db	803376	TTATTTCAATGACACTGTGCGAACCATTAACCATTTACTCAATGCTGGCTATCAAGGGAT	803435
Oy	846	CATCCATGACAGCTAGGCAATGGGAACCTTCCCTTGACTCAAAATGCTTTGAAA	905
Db	803436	TGTATCTGACGAGATTTGGCAATGGAATGTATATGTCACACTTATGATCGCTTAGAAA	803495
Oy	906	AGCAGCCCAATCAGGCGTAGTCGTGCTCGAAGCTCTAGAGTGGGCAATGTTCCACCAC	965
Db	803496	AGCCCAAAAGATATACGCTGTGTGTGCTGCTTCTCTGTAACCAAGGGTTATACAC	803555
Oy	966	CCAAAGGCTGAAGTGAATGAAGAACTTGGTTTGGGTACAGAGAGTCTCAACC	1025
Db	803556	TGCTACGCTGAAGTGAATGAATGAATGAATGATGCTTTGTAAGCATCGATCTTAAATCC	803615
Oy	1026	TCAAAAAGCAGAGTCTTCTTATGTTAGCCCTCACCAAAACTAGATAGAGAGCGAT	1085
Db	803616	ACAGAAAGCCCGCGTCTTGCAATTTAGCTTAACTCAAACTAAGATCCNAAAGTAAT	803675
Oy	1086	CCAAAAGATCTTCTCCACTATTTAA 1110	
Db	803676	TCACAAATATTTTCGAAGACTCTTAA 803700	
RESULT 5			
AAT96346			
ID	AAT96346	standard; cDNA; 1848 BP.	
XX	AAT96346;		
AC			
DT			
XX	08-APR-1998	(first entry)	
XX			
DE	Chimeric gene containing anti-asparaginase MAb light and heavy chain.		
XX			
KW	Immunoglobulin; Ig; heavy chain; variable region; murine; human;		
KW	asparaginase II; monoclonal antibody; MAb; light chain;		
KW	recombinant chimeric polypeptide; ss.		
OS	Synthetic.		
OS	Chimeric - Homo sapiens.		
OS	Chimeric - Mus sp.		
XX			
PN	US5686579-A.		
PD	11-NOV-1997.		
XX			
PF	23-MAY-1995;	95US-0447422.	
XX			
PR	22-JUN-1993;	93US-0081410.	

PR 21-JUN-1988; 88US-0205748.
PR 31-AUG-1992; 92US-0938505.
PR 23-MAY-1995; 95US-0447422.
XX
PA (HYBR-) HYBRISSENS LTD.
XX
XX
PI Ramjeesingh M, Rothstein A, Shami EY;
XX
DR WPI, 1997-558200/51.
XX
XX Self-protecting chimeric polypeptide comprising biologically active
PI sequence and single-chain antibody sequence - has resistance to e.g.
PI disrupting temperature, presence of proteolytic enzymes, etc.
XX
PS Example 2; Columns 27-30; 29pp; English.

The present sequence is a chimeric gene containing the CDRs for the immunoglobulin (Ig) heavy and light chain variable regions of a murine anti-asparaginase II monoclonal antibody (MAb), and human asparaginase II. The gene was used in the preparation of a novel recombinant chimeric polypeptide comprising a 1st region comprising a biologically active domain and another domain containing an epitope, linked via a polypeptide to a 2nd region including a single chain antibody (SCA) having the light and heavy chains of an antibody variable region which specifically binds the epitope in the 1st region. The chimeric polypeptide assumes a conformation in which the SCA is bound to the epitope of the 1st region and protects its biological activity from deactivation by denaturing temperatures or pH conditions, proteolytic enzymes, oxidising agents or alcohol. The regions of the chimeric polypeptide interact to form a structure analogous to an antibody-antigen complex. A L-asparaginase-SCA fusion protein of the above type has better trypsin resistance than free L-asparaginase.

Sequence 1848 BP; 435 A; 498 C; 517 G; 398 T; 0 other;	
Query Match	25.9%; Score 294; DB 18; Length 1848;
Best Local Similarity	56.3%; Pred. No. 9, 7e-82;
Matches 574; Conservative	0; Mismatches 440; Indels 6; Gaps 1

QY	11	GGTCCGGATCCAGGCGCTCTGTTTGATAGGGTAAACCCGAAGTACATATCCACCA	150
Db	835	GGTGTGGGTGGGGTGGCGCGCGGATCTGATCTTACCCAAATATACCATTTTTCACACC	894
QY	151	GGAGCGACCATTGCTGTGTTCGGGGGAATCTAGCGTCAAGAGTACTCTGCTGGAGCA	210
Db	895	GGCGGGACCATTTGCCGGTGTGTGGTACGCCGACCCAAATATTAACACAGTGGGTAA	954
QY	211	GTCACCGTTGATTAAGCTTCTGCAGCCGTCCCTGCCATTCACGACTTAGCCACATCAAG	270
Db	955	GTTTGCGGTAGAAATCTGGTTAATCGCGCCGCAACTAAAGACATTCGCGAAGTTTAA	1014
QY	271	GGTGAACAGATCTCAAGCATTTGGTCCCAAGAGATAGCGGTGAAGGTGTGCTTAACTA	330
Db	1015	GGCGAGCAGTAGTAGTAATATCGGCTCCAGAGCATGAAGATAAATGTCTGGGTGACATG	1074
QY	331	GCCAAAGCTGTCAATGAGCTCCTCGGCCAAAAGAGACCGGAAGCCGTATATCAACCAT	390
Db	1075	GGCAAAAAAATTAA-----CACCACTGCGATGAAGACGAGCGGCTTCCTCAATTACCCAC	1128
QY	391	GGAACTGACACCATGGAAGAGACCGGTTTCTTCACCACTCAGCGTGAAGAAAGCCAAAA	450
Db	1129	GGTACCGAATCAGATGGAAGAAATCTGTTACTTCTCTGACTACGTTGAAATGGCGAATA	1188
QY	451	CCTGTCCTGCTTGTAGAGCGCATGCGTCCAGGCTCTTCATGAGTGTGATGSCCCATG	510
Db	1189	CCGGTGTGATGTGGCGCAATGCGTCGCTCAGCTATATGAGCGCAACGCTGCATTC	1244
QY	511	AATCTCTAATAAGCGCTGAAATGAGATCAACAAAGCCTATCAATCAAAAGAGGGTG	570
Db	1249	AACCTCTATTAAGCGGTAGTGAACCGAGCTGATTAAGGCTTCGCCCAACCGTGGCTGCTG	1300

Accession	Gene	Strain	Location/Qualifiers
QY 571	ANTIGEN-1	ATCC 29216	ATCC 29216
QY 1309	ANTIGEN-2	ATCC 29216	ATCC 29216
QY 631	ANTIGEN-3	ATCC 29216	ATCC 29216
QY 1369	ANTIGEN-4	ATCC 29216	ATCC 29216
QY 691	ANTIGEN-5	ATCC 29216	ATCC 29216
QY 1429	ANTIGEN-6	ATCC 29216	ATCC 29216
QY 751	ANTIGEN-7	ATCC 29216	ATCC 29216
QY 1489	ANTIGEN-8	ATCC 29216	ATCC 29216
QY 811	ANTIGEN-9	ATCC 29216	ATCC 29216
QY 1549	ANTIGEN-10	ATCC 29216	ATCC 29216
QY 871	ANTIGEN-11	ATCC 29216	ATCC 29216
QY 1609	ANTIGEN-12	ATCC 29216	ATCC 29216
QY 931	ANTIGEN-13	ATCC 29216	ATCC 29216
QY 1669	ANTIGEN-14	ATCC 29216	ATCC 29216
QY 991	ANTIGEN-15	ATCC 29216	ATCC 29216
QY 1729	ANTIGEN-16	ATCC 29216	ATCC 29216
QY 1051	ANTIGEN-17	ATCC 29216	ATCC 29216
QY 1789	ANTIGEN-18	ATCC 29216	ATCC 29216
QY 1848	ANTIGEN-19	ATCC 29216	ATCC 29216
QY 1908	ANTIGEN-20	ATCC 29216	ATCC 29216
QY 1968	ANTIGEN-21	ATCC 29216	ATCC 29216
QY 2028	ANTIGEN-22	ATCC 29216	ATCC 29216
QY 2088	ANTIGEN-23	ATCC 29216	ATCC 29216
QY 2148	ANTIGEN-24	ATCC 29216	ATCC 29216
QY 2208	ANTIGEN-25	ATCC 29216	ATCC 29216
QY 2268	ANTIGEN-26	ATCC 29216	ATCC 29216
QY 2328	ANTIGEN-27	ATCC 29216	ATCC 29216
QY 2388	ANTIGEN-28	ATCC 29216	ATCC 29216
QY 2448	ANTIGEN-29	ATCC 29216	ATCC 29216
QY 2508	ANTIGEN-30	ATCC 29216	ATCC 29216
QY 2568	ANTIGEN-31	ATCC 29216	ATCC 29216
QY 2628	ANTIGEN-32	ATCC 29216	ATCC 29216
QY 2688	ANTIGEN-33	ATCC 29216	ATCC 29216
QY 2748	ANTIGEN-34	ATCC 29216	ATCC 29216
QY 2808	ANTIGEN-35	ATCC 29216	ATCC 29216
QY 2868	ANTIGEN-36	ATCC 29216	ATCC 29216
QY 2928	ANTIGEN-37	ATCC 29216	ATCC 29216
QY 2988	ANTIGEN-38	ATCC 29216	ATCC 29216
QY 3048	ANTIGEN-39	ATCC 29216	ATCC 29216
QY 3108	ANTIGEN-40	ATCC 29216	ATCC 29216
QY 3168	ANTIGEN-41	ATCC 29216	ATCC 29216
QY 3228	ANTIGEN-42	ATCC 29216	ATCC 29216
QY 3288	ANTIGEN-43	ATCC 29216	ATCC 29216
QY 3348	ANTIGEN-44	ATCC 29216	ATCC 29216
QY 3408	ANTIGEN-45	ATCC 29216	ATCC 29216
QY 3468	ANTIGEN-46	ATCC 29216	ATCC 29216
QY 3528	ANTIGEN-47	ATCC 29216	ATCC 29216
QY 3588	ANTIGEN-48	ATCC 29216	ATCC 29216
QY 3648	ANTIGEN-49	ATCC 29216	ATCC 29216
QY 3708	ANTIGEN-50	ATCC 29216	ATCC 29216
QY 3768	ANTIGEN-51	ATCC 29216	ATCC 29216
QY 3828	ANTIGEN-52	ATCC 29216	ATCC 29216
QY 3888	ANTIGEN-53	ATCC 29216	ATCC 29216
QY 3948	ANTIGEN-54	ATCC 29216	ATCC 29216
QY 4008	ANTIGEN-55	ATCC 29216	ATCC 29216
QY 4068	ANTIGEN-56	ATCC 29216	ATCC 29216
QY 4128	ANTIGEN-57	ATCC 29216	ATCC 29216
QY 4188	ANTIGEN-58	ATCC 29216	ATCC 29216
QY 4248	ANTIGEN-59	ATCC 29216	ATCC 29216
QY 4308	ANTIGEN-60	ATCC 29216	ATCC 29216
QY 4368	ANTIGEN-61	ATCC 29216	ATCC 29216
QY 4428	ANTIGEN-62	ATCC 29216	ATCC 29216
QY 4488	ANTIGEN-63	ATCC 29216	ATCC 29216
QY 4548	ANTIGEN-64	ATCC 29216	ATCC 29216
QY 4608	ANTIGEN-65	ATCC 29216	ATCC 29216
QY 4668	ANTIGEN-66	ATCC 29216	ATCC 29216
QY 4728	ANTIGEN-67	ATCC 29216	ATCC 29216
QY 4788	ANTIGEN-68	ATCC 29216	ATCC 29216
QY 4848	ANTIGEN-69	ATCC 29216	ATCC 29216
QY 4908	ANTIGEN-70	ATCC 29216	ATCC 29216
QY 4968	ANTIGEN-71	ATCC 29216	ATCC 29216
QY 5028	ANTIGEN-72	ATCC 29216	ATCC 29216
QY 5088	ANTIGEN-73	ATCC 29216	ATCC 29216
QY 5148	ANTIGEN-74	ATCC 29216	ATCC 292

Query Match	Best Local Similarity	Score	DB	Length
Matches	548; Conservative	0; Mismatches	428; Indels	12; Gaps
126	ACCCCAAGTACTATCTTCTAGCCACAGAGAGGACACATCGCTGGTTCCGGGGAATCTAGCCT	185		
157	ACCACCACTTGGCTTTACTGCGACAGAGGGGAGAGATGAGAGAGAGTGGCCGAGCCGAG	216		
186	CAGAGTACTACTCTGCTGAGAGACATCACCCTTGATAGCTCTTTCGACGCGCTCCAG	245		
217	TTGGGGTACTTAAAGATGGTGAGTGGGCAATCAAGAGCTTTGAAGGTATTCCTTAG	276		
246	CATCAACGACATAGCCACCATCAAGGTTAACACAGATCTCAAGCATTTGGCTCCCAAGAT	305		
277	TCTTTACAGACATCGCTCCGATTCACAGGAGAGAGATTTCTAACATCGCTCACACACAT	336		
306	GAGGGTAAAGTGTGGCTTAACTAGCCAGAGGTTGATCAATGAGCTCTCGCCCAAAAAGA	365		
337	GATTAAGAGATAGTGGTTCAAGCTCGCCAAAGCTGGCCAAAGATTGCTAGATGATAGCCG	396		
366	GACCGAAGCGGTGATCATACCCATGGAATCTAGACACATGGAAGAGACCGTTTCTTCT	425		
397	TATTCAGAGCGGTGATCATACCCATGGAAGAGACACATTTAGAGAGAGCGCTATTTTTT	456		
426	CAACCTCAGCGTGAAGAGCAAAAACCTGTGCTTGGCTTGAAGCGGCATCGTCCAGGCTC	485		
457	AACTTACTTATAGCTCCACAAAACCGGTCGTGCTGGTGGAGAGCGATCGTAAATGCTGC	516		
486	TTCCATGAGTGTGATGAGCCCATGAAATCTCTAATAAGCGGTGAATGTAGCGATCAACAA	545		
517	TTCTTTAGAGCGGTGATGAGCGCTTAAATTAATGATGCTGAGCGTGCCTCAATGA	576		
546	AGCCTCTACTAACAAGAGGTGATGTGATGAAACGATGAGATTCACGCGCCCAAGAGA	605		
577	AAAAAGTGGATTAAGAGCGGTGATGTGATGAGCAATTAATTTTGAAGCGCTAGAGA	636		
606	AGCGACCAAGCTCAACACACCGGACGATCAATGATCTTGGTCCCAACACAGTAAAT	665		
637	AGTATTAAGAGCGACACACACCGGATCTTCAAGCTTAAATAGGCGGCAT	696		
666	CGGACATGATTTATGCGAAGTGAATTTTCACTAATCGTTCGATCGACACCT	725		
697	AGGAGCGCTGATTTATGCGAAGTGAATTTTCACTAATCGTTCGATCGACACCT	756		
726	TGCAAGTATGATTTATGCAAAATGAA---GACCTCCCAAGTGCATCTCTTA	782		
757	AGAGAGCAATTTTCCCTTTCAACACTAAAAACCCCTTGCCTTAAGTGAATATATTA	816		
783	CGCTACCCGATGATGATGATTTTATGCAATGACACCGCTTCGAGGAGAGCAACAAAG	842		
817	CACGATGCTGATGATGATGATTTTATGCAATGACACCGCTTCGAGGAGAGCAACAAAG	876		
843	AATCATCATGATGATGATGATTTTATGCAATGACACCGCTTCGAGGAGAGCAACAAAG	902		
877	CGTGTGATGATGATGATGATTTTATGCAATGACACCGCTTCGAGGAGAGCAACAAAG	936		

QY 903 AAAAGCAGCCAAATCAGGCTAGTCGCTCCAGCTTAGAGTGGGAGTGGTCCAC 962
 DB 937 AGAAGCAGCCAAATGAGGCTAGTCGCTTAGAGTGGGAGTGGTCCAC 966
 QY 963 CACCAAGAGGCTAGTGGTATGATGAAAGCTGTTTGTGCTACAGAGTCTCAA 1022
 DB 997 TA-----CTTCAGGCGGATGATGATGACAGGCTTATCAGAGGCAATTTAA 1047
 QY 1023 CCCTCAAAAAGCCAGAGTGTCTTATGTTAGCCCTCACCAGAACTAGTATAGAGGC 1082
 DB 1048 CCCCCAAAAGCTAGGCTGCTTTACAACTGCTTAACTAAACAAATATATAAGAAA 1107
 QY 1083 GATCCAAAGATCTTCTCCAGCTTTAA 1110
 DB 1108 ATCCAAAGAAATGTTTGAAGATATGA 1135

RESULT 7
 ID AAN70557 standard; DNA; 2837 BP.
 AC AAN70557;
 XX 25-MAR-2003 (updated)
 DT 07-AUG-1991 (first entry)
 XX
 DE Sequence encoding a protein with L-asparaginase (LA) activity in
 DE PASN 30 or PASN 32.
 XX
 KW Enzyme; cancer therapy; leukemia therapy; ss.
 OS Erwinia chrysanthemi.
 XX
 FH Key Location/Qualifiers
 FT misc_feature /tag= a
 FT /label= pUC9 linker
 FT sig_peptide /tag= b
 FT mat_peptide /tag= c
 FT /note= "claimed"
 FT terminator /tag= d
 FT
 XX
 EP211639-A.
 PD 25-FEB-1987.
 XX
 PF 04-AUG-1986; 86EP-0305984.
 PR 06-AUG-1985; 85GB-0019753.
 XX
 PA (PUBL-) PUBLIC HEALTH LAB SERVICE BOARD.
 PI Atkinson A, Minton NP, Gilbert HU;
 DR WPI: 1987-051784/08.
 DR P-PSDB; AAP70348.
 XX
 PT New recombinant plasmids coding for L-asparaginase - esp. from
 PT Erwinia carotovora, useful for treating malignancies, and new
 PT transformed hosts
 XX
 PS Disclosure; Fig 4; 35pp; English.
 XX
 CC When E. carotovora SCRI 193 conty. plasmid PASN 32 was cultured, the
 CC cell-free culture broth contained 11.19 LA units/ml with specific
 CC activity 53.34 units/mg of protein. These figures compare with 10.87
 CC LA units/ml and 14.97 units/mg of protein for the currently used
 CC prodn. strain E. chrysanthemi NCPPB 1066.
 CC (updated on 25-MAR-2003 to correct PA field.)

XX
 SQ Sequence 2837 BP; 627 A; 665 C; 867 G; 678 T; 0 other;
 Query Match 19.7%; Score 223.2; DB 8; Length 2837;
 Best Local Similarity 53.3%; Pred. No. 2.7e-59;
 Matches 527; Conservative 0; Mismatches 443; Indels 18; Gaps 2;

QY 127 CCCCAGTACTATCCCTAGCCACAGAGGACCAATCGCTGGTGGGGAAATC---TAGC 183
 DB 736 CCAATATCCGTATACCTGCGGACCGGCGGTACAAATGCGCGCTCAGCGGAGTACC 795
 QY 184 GTCAAGTACTACTCTGTGAGAGTCAACCGTTGATTAAGCTTCTTCAGCCGCTCC 243
 DB 796 CAACACAGGTTACAGAGTGGCGGCTGGCGTGAATGATTAACGCTGTGCT 855
 QY 244 GCAATCAAGCAGCTAGCCACCATCAAGGTTGAAGATGATCAAGCTAGGCTCCCAAG 303
 DB 856 GAGTGAAGAACTAGCTTAATGTGAAGGGGAGAGCATCTCCAAATAGGCCAGGAAAC 915
 QY 304 ATGACGGGTAAAGTGTGCTTAACTAGCCAGCGTGTCAATGAGCTCTCGCCCAAAA 363
 DB 916 ATGACGGGTAAAGTGTGCTTAACTAGCCAGCGTGTCAATGAGCTCTCGCCAGAT 975
 QY 364 GAGACCGAAGCCGTATCATCAACCATGGAATGACACCATGGAAGAGAGAGCGCTTCT 423
 DB 976 GATGTGATGTGTGTGTATCAACCAAGGACGACAGCGTGAAGAGTGGCTTACTTT 1035
 QY 424 CTCACCTCAGGCTGAAAGCCAAACCTGTCTGCTGTGAAGGCGCATGCTCCAGGC 483
 DB 1036 CTTCATCTGACGCTGAAAGAGTACCAAGCGATGTTGTGTGCGAGCATGTGCGGGA 1095
 QY 484 TCTTCATGAGTGTGATGAGGCGCCCATGAATCTCTATACCGCTGATATACGATCAAC 543
 DB 1096 ACGGCATCAGCTGACGCGCCGATGAACCTGCTGAAGAGCGTACGCTGCGGTGAC 1155
 QY 544 AAAGCTCTACTATACAAAGAGTGGTATGATGAAGATGATGATGATGATGATGATGAT 603
 DB 1156 AAACAGTCTCGCGGTGCGCGCGGTGATGATGATGATGATGATGATGATGATGATGAT 1215
 QY 604 GAAGGACCAAGCTACCAACCAACGAGTCAATGATGATGATGATGATGATGATGATGAT 663
 DB 1216 TACATCCCAAGCAAGCAAGCCTCTACGATGATGATGATGATGATGATGATGATGAT 1275
 QY 664 ATGCGCACGCTATATATGCAAGAGTCAATGATGATGATGATGATGATGATGATGATGAT 723
 DB 1276 CTGGGCGTCAATATGTAACGCAATTTACTACCAAAACGATGATGATGATGATGATGAT 1335
 QY 724 CTTCGAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 783
 DB 1336 ACCCGGTCTGTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1395
 QY 784 GCTACCCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 843
 DB 1396 GGTATACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1455
 QY 844 ATGATCAGTACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 903
 DB 1456 ATGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1515
 QY 904 AAAGCAGCCAAATCAGGCTAGTCGCTCCAGAGCTCTAGAGTGGGAGTGGTCCACC 963
 DB 1516 AAGGCGATGAGAAAGGCGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 1571
 QY 964 ACCCAAGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1620
 DB 1572 -----TGTGCGCGCGGATGAAGAGCTGCGAGTGTGTTGTGACTCTTAAAC 1620
 QY 1024 CTTCAAAAAGCAGAGTGTCTTATGTTAGCCCTCACCAGAACTAGTATAGAGAGGC 1083
 DB 1621 CCGGCAATGCGGCAATCTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1680
 QY 1084 ATCCAAAGATCTTCTCCAGCTTTAA 1111

Db 1681 ATTCAAGACTATTCCATCTATTGAT 1708

RESULT 8
AA068438
ID AA068438 standard; DNA; 1014 BP.
XX
AC AA068438;
XX
DT 25-MAR-2003 (updated)
DT 12-JAN-1995 (first entry)
XX
DE Pseudomonas glutaminase gene.
XX
KW Glutaminase; antiviral; virucide; anticancer; cancer therapy;
KW HIV virus; gene therapy; Escherichia coli; ds.
XX
OS Pseudomonas sp.
XX
PN MO9413817-A1.
XX
PD 23-JUN-1994.
XX
PF 04-DEC-1992; 92WO-US10421.
XX
PR 04-DEC-1992; 92WO-US10421.
PR 04-DEC-1992; 92AU-0032358.
XX
PA (MEME-) ME MEDICAL ENZYMES AG.
XX
PI Freeman AG, Macallister TW, Roberts J, Sethuraman N;
XX
DR WPI: 1994-217891/26.
DR P-PSDB; AAR59739.
XX
PT Recombinant glutaminase derived from Pseudomonas 7A - expressed
PT in E. coli to increase yield and avoid Pseudomonas endotoxins for
XX antiviral and anticancer therapy
XX
PS Disclosure: Page 33; 60pp; English.
XX
CC Chromosomal DNA from Pseudomonas sp. 7A (ATCC 29538) was used to
CC construct a genomic library in Escherichia coli LB392. Screening
CC with mixed oligonucleotide probes was used to isolate a glutaminase-
CC encoding clone. This was sequenced using the primers given in
CC AA068439-47. The gene can be used to manufacture recombinant
CC glutaminase, free of Pseudomonas exotoxin, for use in e.g. HIV and
CC cancer therapy. The gene may also be used in gene therapy protocols.
CC (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 1014 BP; 226 A; 324 C; 318 G; 146 T; 0 other;

Query Match 19.0%; Score 215.8; DB 15; Length 1014;
Best Local Similarity 52.6%; Pred. No. 3.3e-57;
Matches 519; Conservative 0; Mismatches 462; Indels 6; Gaps 2;

QY 128 CCCAAGTACTATCTCCAGCAGAGCACCATTGCGTGGGGAATCTAGCTCA 187
DB 29 CCAACGTGTGATCTGCGCCACCGCGGCACATGCGCGGCGCCAGCGCGCA 88
QY 188 AGAG--TAGTACTCTGTGGAGCAGTCACTGATAGCTTCTTGAAGCGCTCCCTG 244
DB 89 ACAGCGCCACCTACAGAGCTGCGCAAGGTGGCTGCGAAGCTATTTGCGCGCTGCGG 148
QY 245 CCATTCAGAGACCTAGCACCATCAAGCTGAGCAGATCTAAGCATTTGGCTCCCAAGAGA 304
DB 149 AGCTGGCGACCTGCGCAATGTGCGCGGCGAGTGTGATGCAATGCGCTCCGAAAGCA 208
QY 305 TGAAGGTAAGTGTGCTTAAGTACGAGCGTGTCAATGAGCTCTGCGCCAAAG 364
DB 209 TCAACCAAGCAGACCTCTCAAGTGTGCAAGCAGCGTGGCGGAGCTGGCCGACGCAATG 268
QY 365 AGACCGAAGCGGTGATCATCAACCATGAACTGACACCATGGAAGAGACCGCTTTCTTCC 424

Db 269 ACCTCATGCGATCTGATCATACCATGAGCAGCAGACCTGTGAAGAAACCGCTACTTTT 328
QY 425 TCACCTCAGCGGTGAAGCAAAACCTGTCTGCTTGTAGCGCCATGCTCGAGGCT 484
DB 329 TGAACCTCGTGAAGAAAGACGACGACAGCCGATGCTGTGTGCTTGCATGCGCCCGCA 388
QY 485 CTTCATGAGTGTGATGCGCCCATGATCTCTATTAAGCGCGTGAATGTAGCGATCAAGA 544
DB 389 CCGCATGTCCGCGAGCGCATGCTCAACCTGTACAGCGCGTGGCGCAGCAGACA 448
QY 545 AAGCCTTACTAACAAGAGAGTGTGATTTGATGAAGATGATTCACGCGCCAGAG 604
DB 449 AGGACTCGCGCGCAGAGGCGTGTGTGATGACATGAACAGAGATGCTCGGCGGTG 508
QY 605 AAGCAGCAAGCTCAACACCGCGCATGATGATTTGCTTCCGCCAACAGAGTAA 664
DB 509 ACGTGAGCAAGTCGATCAACATCAAGACCGAAGCTTC--AAGACGCGCTGGGCGCCG 565
QY 665 TCGGACAGTCTATTTATGGAAGTGAAGTATTTCACTCAATCCGTGACCTCACACC 724
DB 566 TGGCATGTGTGTGAAGGCAAGTGTGCTGCTCGCGCGGCGCAAGCGCCACACCG 625
QY 725 TTGCAAGTGAAGTTGATATTAGCAAAATCGAAGAACTCCCGAGTGTATTTCTTACG 784
DB 626 TCAACTCGAGTTGACATCAAGCAGATGACAGAGCTGCGCAGGTGACATCGCTACA 685
QY 785 CTCACCCGATGATGATGATTTTATGATATGCAAGCCCTTACGAGGAGCCAAAGCAA 844
DB 686 GCTATGGAAGCTCACCGACGCGCTTCAAGGCGCTTGCAAGAGCGCGCAAGGCGCG 745
QY 845 TCATCATGAGCAGATGAGCAATGGAACCCCTTCCCTTGTACTCAAAATGCTTGA 904
DB 746 TGATCCATGCGCGCAGACCGCATGCTGCTGCTGCGGCTGCTGCGCGCTGCGAG 805
QY 905 AAGCAGCAATCAAGCGTGTGCTGCTGCAAGCTTGAAGTGGCAGAGGTTTCCACCA 964
DB 806 AGCTCGCAGAGAACCGCTGCAATGCTGCTGCAAGTCAAGAGGCGGTTTCTGTC 865
QY 965 CCCAAGAGCTGAATGATGATTAAGAACTGTTGTGTGCTACAGAGTCTCAAC 1024
DB 866 TCGCTAAGCGCGAGAGCGCGCGCAGCAAGACGATGCTGTGCGCCAGCTGCAAC 925
QY 1025 CTCAAAAGCAGAGTGTCTTATTTAGTGGCTTCCACCAAACTGATGATGAGAGGCGA 1084
DB 926 CGCAGAAAGCGCGATCTTGGGATGTGGCAATGACCAAGACCGAGCAGCAAGAGC 985
QY 1085 TCAGAAAGATCTTCTCCACCTATTAT 1111
DB 986 TGCAGCGCATTTTCTGGGAATGATGAT 1012

RESULT 9
AB280801
ID AB280801 standard; DNA; 1174 BP.
XX
XX AB280801;
XX
DT 30-MAY-2003 (first entry)
DT
XX
DE Erwinia carotovora L-asparaginase gene.
KW gene; ds; asparaginase; cytosolic; blood; acute lymphoblastic leukemia.
XX
OS Erwinia carotovora.
XX
FH Key
FT CDS
FT Location/Qualifiers
FT 48..1094
FT /*tag= "a"
FT /product= "L-asparaginase"
FT /EC number= 3.5.1.1
FT 48..72
FT primer_bind
FT /*tag= b

CC Helicobacter pylori selected from 421 complexes given in the
 CC specification. The complex of protein-protein interactions are useful
 CC for screening for agents which modulate the interaction of proteins.
 CC Modulating compounds which binds to a targeted bacterial protein may be
 CC used for treating or preventing ulcers in a human or animal. This
 CC sequence encodes a selected interacting domain (SID), identified via
 CC protein-protein interactions.
 CC
 SQ Sequence 345 BP; 90 A; 69 C; 104 G; 82 T; 0 other;

Query Match 9.6%; Score 108.8; DB 24; Length 345;
 Best Local Similarity 57.7%; Pred. No. 1.1e-23;
 Matches 194; Conservative 0; Mismatches 142; Indels 0; Gaps 0;

QY 142 CTAGCCACAGGAGGACCATCGCTGTTCCGGGAGATCTAGCGTCAAGATGAGTACTCT 201
 Db 2 CTGGCAGACAGGAGGAGCATTCAGAGGAGTGTGCGAGCCGGAGTTGGTAGTTATAG 61
 OY 202 GCTGAGCAGCTACACCGTTGATACCTTCTTGACCCGCTCCCTGCCATCAGACCTAGCC 261
 Db 62 AGTGGTAGTGGGATCAAGACCTTTTGAAGGCTATCCCTAGCTTAACAGACTGCT 121
 OY 262 ACCATCAAGGAGTGAACAGATCTCAAGATTTGCTCCCAAGAGATGAGCGTAAAGTGG 321
 Db 122 CGCATTCAGAGGAGGAGATTTCTAACATGCGCTCACAAAGCATGAATGAAGAGTATGG 181
 OY 322 CTTAATACGCAAGCGTGTGATGAGCTCTCGCCCAAAAGACCGAGCCGATGC 381
 Db 182 TTCAGAGCTCCCAAAACCTGCCAAGAAATGCTAGATGATAGCCGATTTCAAGAGTGGTC 241
 OY 382 ATCACCATGGAAGTCAACATGAGAGACCGCTTCTTCCCAACCTCAGCGTAA 441
 Db 242 ATCAGCATGCGACGACGACCTTTAGAGAGAGAGCGCGATTTTAAACTAGTTTACGC 301
 OY 442 AGCCAAAACCTGTCTGCTGTAGAGCGCCATGCGT 477
 Db 302 TCCACAAAACCGGTGCTGTGGAGCGATGCGT 337

RESULT 11
 ID AAS21124 standard; cDNA; 543 BP.
 AC AAS21124;
 XX
 DT 21-MAY-2002 (first entry)
 XX
 DE cDNA encoding Tm 3.9/His tag fusion protein minus signal peptide.
 KW Anti-freeze peptide; Tm 3.9; yellow mealworm beetle; TRP; cryoinjury;
 KW thermal hysteresis protein; cryosurgery; cold climatization; ss;
 KW hypothermic cell preservation; de-icing formulation; transgenic plant;
 KW transgenic animal; His tag.
 XX
 OS Tenebrio molitor.
 XX
 FH Key Location/Qualifiers
 FT CDS 64..513
 FT /tag= a
 FT /product= "Tm 3.9 clone protein"
 FT misc-feature 64..165
 FT /tag= b
 FT /note= "Nucleotides encoding His tag"
 FT mat_peptide 166..510
 FT /tag= c
 XX
 PN MO200194378-A1.
 XX
 PD 13-DEC-2001.
 XX
 PF 07-JUN-2001; 2001MO-US18532.
 XX
 PR 08-JUN-2000; 2000US-210446P.
 XX

XX (UNIV) UNIV NEW YORK STATE RES FOUND.
 PA (HORN/) HORNWATH K L.
 PA (MYER/) MYERS K L.
 PA (EAST/) EASTON C M.
 XX
 PI Horwath KL, Myers KL, Easton CM;
 XX
 DR WPI; 2002-090137/12.
 DR P-PSDB; AAU10053.
 XX
 PT New cDNA polynucleotide encoding a thermal hysteresis protein which is
 PT a Type III anti-freeze protein derived from the Tenebrionidea
 PT Superfamily, useful for providing antifreeze protection to improve the
 PT quality of food -
 XX
 PS Claim 4; Page 345; 364pp; English.
 XX
 CC This invention relates to a cDNA polynucleotide comprising a nucleotide
 CC sequence encoding a thermal hysteresis protein (TRP) which is a Type III
 CC anti-freeze protein derived from the Tenebrionidea Superfamily. Anti-
 CC freeze proteins lower the freezing point of a solution without affecting
 CC the melting point of the solution. An activated anti-freeze protein may
 CC be incorporated into plant, produce or fish in an amount sufficient to
 CC provide antifreeze protection or in a region of a target tissue to
 CC provide antifreeze protection to limit tumor cell or target tissue
 CC cryoinjury during cryosurgery. The proteins of the invention may also be
 CC used in hypothermic solutions or bathing media to reduce cold damage in
 CC order to provide cryogenic or hypothermic preservation of cells and
 CC tissues. The proteins may be used as de-icing formulations or used on
 CC surfaces to reduce existing ice buildup or abate the formation of ice
 CC buildup on surfaces such as a road, aircraft, household products,
 CC machinery and plant surfaces or as a food product to improve the quality
 CC of food by abating freezing of solutions, freezer burn, or degradation
 CC due to cold storage. The polynucleotides for the activated protein can
 CC be used to create transgenic or gene-modified plants, crops, fish, or
 CC animals having greater tolerance to cold climatization. The Tm 12.86
 CC antibody/antisera which is also used as a screening device to screen
 CC cDNA libraries in an expression system, including cross-species cDNA
 CC libraries to identify homologous sequences in other species. The
 CC present sequence represents a cDNA encoding the Tm 3.9 clone/His
 CC tag fusion protein minus the signal peptide of the invention. This cDNA
 CC was created to facilitate purification and to try to enhance the anti-
 CC freeze activity of the recombinant protein.
 CC
 SQ Sequence 543 BP; 164 A; 133 C; 144 G; 102 T; 0 other;

Query Match 9.1%; Score 103.6; DB 24; Length 543;
 Best Local Similarity 88.9%; Pred. No. 5.9e-22;
 Matches 112; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 1 ATGGGACAGCAGCATATCATCATCATCATTAGAGGCGGCTGTGGCGGCGGACCAT 60
 Db 64 ATGGGACAGCAGCATATCATCATCATCATCAGCAGGCGGCTGTGGCGGCGGACCAT 123
 OY 61 ATGGGACAGCAGCTGTGTGAGCAGCAAAATGGTGGGATCCAGCCCTCTGTTTGATG 120
 Db 124 ATGGGACAGCAGCTGTGTGAGCAGCAAAATGGTGGGATCCCTGATCAACGATGAACAGATA 183
 OY 121 GCTAA 126
 Db 184 CAGAAA 189

RESULT 12
 ID AAS21116 standard; cDNA; 543 BP.
 AC AAS21116;
 XX
 DT 21-MAY-2002 (first entry)
 XX
 DE cDNA encoding Tm 2.2 clone/His tag fusion protein minus signal sequence.
 XX

be incorporated into plant, produce or fish in an amount sufficient to provide antifreeze protection or in a region of a target tissue to provide antifreeze protection to limit tumour cell or target tissue cryoinjury during cryosurgery. The proteins of the invention may also be used in hypothermic solutions or bathing media to reduce cold damage in tissues. The proteins may be used as de-icing formulations or used on surfaces to reduce existing ice buildup or abate the formation of ice buildup on surfaces such as a road, aircraft, household products, machinery and plant surfaces or as a food product to improve the quality of food by abating freezing of solutions, freezer burn, or degradation due to cold storage. The polynucleotides for the activated protein can be used to create transgenic or gene-modified plants, crops, fish, or animals having greater tolerance to cold climatisation. The Tm 12.86 antibody/antisera which is also used as a screening device to screen cDNA libraries in an expression system, including cross-species cDNA libraries to identify homologous sequences in other species. The present sequence represents a cDNA encoding the Tm 2.3 clone/his tag fusion protein of the invention. This cDNA was created to facilitate purification and to try to enhance the anti-freeze activity of the recombinant Tm 2.3 protein.

Sequence 543 BP; 164 A; 133 C; 143 G; 103 T; 0 other;

Query Match 9.0%; Score 102; DB 24; Length 543;

Best Local Similarity 88.1%; Pred. No. 1.9e-21;

Matches 111; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

```

OY 1 ATGGGACAGCAGCATCATCATCATCATATAGACAGCGCTGTGCGCGGAGCCAT 60
    |||||
DB 64 ATGGGACAGCAGCATCATCATCATCATATAGACAGCGCTGTGCGCGGAGCCAT 123
    |||||
OY 61 ATGGCTACATGACTGTGTGACAGCAATGGGTGCGGATCCAGCGCTCTGTTGTATG 120
    |||||
DB 124 ATGGCTACATGACTGTGTGACAGCAATGGGTGCGGATCCAGCGGACAGATATA 183
    |||||
OY 121 GCTAAA 126
    |||
DB 184 CAGAAA 189

```

RESULT 14

AAS21122 standard; cDNA: 543 BP.

AC AAS21122;

DT 21-MAY-2002 (first entry)

DE cDNA encoding Tm 3.4/his tag fusion protein minus signal sequence.

XX Anti-freeze peptide; Tm 2.3; yellow mealworm beetle; THP; cryoinjury;
 XX thermal hysteresis protein; cryosurgery; cold climatisation; ss;
 KW hypothermic cell preservation; de-icing formulation; transgenic plant;
 KW transgenic animal; His tag.

OS Tenebrio mollitor.

XX Key Location/Qualifiers

XX CDS 64..513

FT /*tag= a /product= "Tm 3.4 clone protein"

FT misc_feature 76..93

FT /*tag= b /note= "Nucleotides encoding His tag"

FT mat_peptide 166..510

XX /*tag= c

XX WO200194378-A1.

XX 13-DEC-2001.

XX 07-JUN-2001; 2001WO-US18532.

XX 08-JUN-2000; 2000US-210446P.
 PR (UNY) UNIV NEW YORK STATE RES FOUND.
 XX (HORN/) HORWATH K L.
 PA (MAYER/) MYERS K L.
 XX (EAST/) EASTON C M.
 PI Horwath KL, Myers KL, Easton CM;
 DR WPI; 2002-090137/12.
 DR P-PSDB; AAU10051.
 XX
 PT New cDNA polynucleotide encoding a thermal hysteresis protein which is
 PT a Type III anti-freeze protein derived from the Tenebrionidae
 PT Superfamily, useful for providing antifreeze protection to improve the
 PT quality of food -
 XX
 PS Claim 4; Page 340; 364pp; English.

This invention relates to a cDNA polynucleotide comprising a nucleotide sequence encoding a thermal hysteresis protein (THP) which is a Type III anti-freeze protein derived from the Tenebrionidae Superfamily. Anti-freeze proteins lower the freezing point of a solution without affecting the melting point of the solution. An activated anti-freeze protein may be incorporated into plant, produce or fish in an amount sufficient to provide antifreeze protection or in a region of a target tissue to provide antifreeze protein to limit tumour cell or target tissue cryoinjury during cryosurgery. The proteins of the invention may also be used in hypothermic solutions or bathing media to reduce cold damage in tissues. The proteins may be used as de-icing formulations or used on surfaces to reduce existing ice buildup or abate the formation of ice buildup on surfaces such as a road, aircraft, household products, machinery and plant surfaces or as a food product to improve the quality of food by abating freezing of solutions, freezer burn, or degradation due to cold storage. The polynucleotides for the activated protein can be used to create transgenic or gene-modified plants, crops, fish, or animals having greater tolerance to cold climatisation. The Tm 12.86 antibody/antisera which is also used as a screening device to screen cDNA libraries in an expression system, including cross-species cDNA libraries to identify homologous sequences in other species. The present sequence represents a cDNA encoding the Tm 3.4 clone/his tag fusion protein of the invention. This cDNA was created to facilitate purification and to try to enhance the anti-freeze activity of the recombinant Tm 3.4 protein.

Sequence 543 BP; 163 A; 134 C; 144 G; 102 T; 0 other;

Query Match 9.0%; Score 102; DB 24; Length 543;

Best Local Similarity 88.1%; Pred. No. 1.9e-21;

Matches 111; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

```

OY 1 ATGGGACAGCAGCATCATCATCATCATATAGACAGCGCTGTGCGCGGAGCCAT 60
    |||||
DB 64 ATGGGACAGCAGCATCATCATCATCATATAGACAGCGCTGTGCGCGGAGCCAT 123
    |||||
OY 61 ATGGCTACATGACTGTGTGACAGCAATGGGTGCGGATCCAGCGCTCTGTTGTATG 120
    |||||
DB 124 ATGGCTACATGACTGTGTGACAGCAATGGGTGCGGATCCAGCGGACAGATATA 183
    |||||
OY 121 GCTAAA 126
    |||
DB 184 CAGAAA 189

```

RESULT 15

AAS21126 standard; cDNA: 543 BP.

AC AAS21126;

DT 21-MAY-2002 (first entry)

Db	802956	TACTGATGGAATTTGTCTATTACCCATGGTACAGATACCATGGAAGAGACGGCTTAATTTCT	803015
QY	426	CAACCTCAGCGTGAAGAAAGCCAAAAACCTGTCTGTGTAGGGCCATGCGTCCAGGCTC	485
Db	803016	AGATTTAACCGGTAAATGTGAAAAACCGGTTGTTCTCGTTGGGGCAATGCGCTCCGCAAC	803075
QY	486	TTCCATGAGTGTCTATGGCCCCAGTAATCTGTATAAGCCGGTGAATGTAGGATCAACAA	545
Db	803076	AGAAAAAGTGTCTATGGCCCATTTAAATCTTTACATCTGTCTGTTTGGCAGACAGACAA	803135
QY	546	AGCCTCTACTTAACAAAGAGTGTGATTTGTATGATGAACGATGAGATTCAAGCCGCCAGAGA	605
Db	803136	AAATTCAAAGT	803195
QY	606	AGGAGCCAGCTCAACACCCACCGCAGTCAATGCTATTGTGTTGCCACACAGGTAAAT	665
Db	803196	TGTAAACAAAAACAGTACGACCGCAGTCAAAACGTTCCATTTCACCAATTAATGTGTTCT	803255
QY	666	CGGACAGTCTATTATTTGCGAAAGTGTGATTTTACTCAATCCGTTTGCACCTCACACCT	725
Db	803256	AGGCTATTTCATTAACACCAAAAGTGTGATGATGAAGTGTGTGTGTGTGTGTGTGTGTGT	803315
QY	726	TGCAAGTAGTTGATTAATTAGCAAAATGGAAGACTCCCGCAGTGTGATTTCTTTACGC	785
Db	803316	CAACACTCGGTTTAACGTAGAAAAATTAATGATGCTTACCAAGTGGGGATTATTATGC	803375
QY	786	TCACCCGATGTACTGTATGTTTGTGCAATGACGCTTCAGGCGAGAGCCAAAGGAAT	845
Db	803376	TTATTCAATGACCTGTTCGAAACCTTAACGCAATTACTCAATGCTGTGTATCAAGGAT	803435
QY	846	CATCCATGACGAGCTAGGCGAATGGAAACCTTTCCCTTTGACTCAAAATGCTCTTGA	905
Db	803436	TGTATCTGAGAGAGTTGGCAATGGAATGTATGCTGACACTTGTATCGCTTAAGAAA	803495
QY	906	AGCACCACCAATAGCGCTGTGCTGTGCTGCAAGCTTACAGTGGGCGAGTGTCCACAC	965
Db	803496	AGCCCAAAAAAGATGT	803555
QY	966	CCAGAGAGCTGAAGTGTATGATTAAGAACTGTGTTTGTGTGCTACAGAGAGTCTCAACC	1025
Db	803556	TCGTGTGCGCTGAAGTTGATGATAGTAAATGATGCTTTGAGCATVCAGGTCTTAATCC	803615
QY	1026	TCAAAAAGCCAGAGTCTTCTATGTTAGCCCTCACCAAAACATAGATAGAGAGCGAT	1085
Db	803616	ACAGAAAGCCCGGTGTCTCTTGCAATGCTTTAACTCAACTAAGATCCMAAAGTAAT	803675
QY	1086	CCAAAAGATCTTCTCCACCTATTAA	1110
Db	803676	TCACAATAATTTGAGAGACTTTCTAA	803700
RESULT 3			
US-09-643-990A-1			
: Sequence 1, Application US/09643990A			
: Patent No. 6528289			
GENERAL INFORMATION:			
APPLICANT: Robert D. Fleischmann			
Mark D. Adams			
Owen White			
Hamilton O. Smith			
J. Craig Venter			
TITLE OF INVENTION: The Nucleotide sequence of			
the Haemophilus influenzae Rd Genome, Fragments			
thereof, and Uses Thereof			
NUMBER OF SEQUENCES: 1			
CORRESPONDENCE ADDRESS:			
ADDRESSEE: Human Genome Sciences, Inc.			
STREET: 9410 Key West Avenue			
CITY: Rockville,			
STATE: MD			
COUNTRY: USA			
ZIP: 20850			

Query Match	26.2%	Score 296.4	DB 4	Length 1830121
Best Local Similarity	57.1%	Pred. 2.9e-86		
Matches	562	Conservative	0	Mismatches 417; Indels 6; Gaps 1
<p>COMPUTER READABLE FORM:</p> <p>MEDIUM TYPE: 3 1/2 inch diskette</p> <p>COMPUTER: Dell Pentium</p> <p>OPERATING SYSTEM: MS DOS v6.22</p> <p>SOFTWARE: ASCII Text</p> <p>CURRENT APPLICATION DATA:</p> <p>APPLICATION NUMBER: US/09/643,990A</p> <p>FILING DATE: 23-Aug-2000</p> <p>CLASSIFICATION: <Unknown></p> <p>PRIOR APPLICATION DATA:</p> <p>APPLICATION NUMBER: 08/487,429</p> <p>FILING DATE: 1995-06-07</p> <p>APPLICATION NUMBER: 08/426,787</p> <p>FILING DATE: 1995-04-21</p> <p>ATTORNEY/AGENT INFORMATION:</p> <p>NAME: Kenley K. Hoover</p> <p>REGISTRATION NUMBER: 40,302</p> <p>REFERENCE/DOCKET NUMBER: PBI86P1C1</p> <p>TELECOMMUNICATION INFORMATION:</p> <p>TELEPHONE: 301-610-5790</p> <p>TELEFAX: 310-309-8439</p> <p>INFORMATION FOR SEQ ID NO: 1:</p> <p>SEQUENCE CHARACTERISTICS:</p> <p>LENGTH: 1830121 base pairs</p> <p>TYPE: nucleic acid</p> <p>STRANDEDNESS: double</p> <p>TOPOLOGY: linear</p> <p>SEQUENCE DESCRIPTION: SEQ ID NO: 1:</p> <p>US-09-643-990A-1</p>				
126	ACCCCAATGACTATCTAGTACGACACAGAGGAGGACCATCCGTGGTTGCGGGGAATCTAGCGT	185		
802722	ACCAATATTTTCAATCTGGGACAGGCGGTGTACCATTCGACAGAAAGCGGCAAAAGTTCGCT	802781		
186	CAAGATGACTACTCTGCTGCGAGCAGTCAACGTTGATTAAGCTCTTGTGAGCGCTGCCCTGC	245		
802782	AAATTCCTCGTATTAAGCTGGCAATTAAGTATGATCTTAATTAAGAGGCTGTACCCAGA	802841		
246	CATCAAGCAGCTAGCCACCATCAAGGATGAGACAGATCTCAAGCATTTGCTCCACAGAGAT	305		
802842	AATGAAAAATATGGCCACATTAAGAGTGAAGCAAAATTTGAAAAATAGGTTTCAAGACAT	802901		
306	GACGGTAAAGTGTGGCTTAAACATGACCAAGCGTGTCAATGAGTCTCTCGCCCAAAAAAGA	365		
802902	GATGTCGAAAGTCTGGCTTAAACCTGGCAAAAGCCATCAATGCTC-----AATGTAAAG	802955		
366	GACCGAAGCGTATCATCACCATGGAATGACATGACATGAGAAAGACGCGCTTCTCTCT	425		
802956	TACTGATGATTTGTTCATTAACCATGCTACAGATACCATGGAAGAGACGGCTTAATTTCT	803015		
426	CAACCTCAGCGTGAAGCAAAACCTGTCTGCTTTGTAGCGCCATCGCTCCAGGCTC	485		
803016	AGATTTAACCTGTAATGTGAAGAAACCGGTTGTCTTCCTGGGCAATCGCTCTGCAAC	803075		
486	TTTCATGATGCTGCGATGGCCCATGAATGCTTAAAGCGCGTAATGTAGCATCAACAA	545		
803076	AGAAAAAGTGTGATGGCCCATTAATATCTTTTCAATGCTGTCTGTGCGCAGACAGAA	803135		
546	AGCCTTACTAACAAAGAGTGTGATTTGATGAACGATGAGATTACGCGCCAGAGA	605		
803136	AAATCAAGTGTGTGTGTTTGTAGCGCAATGAATTAAGTACTGTAGGTCTCTGCGCA	803195		
606	AGCGACAGCTTAACACGACCGCATGATGATTTGCTGGCCCAACACAGGTAAAT	665		
803196	TGTAAACAAAAACAGTACGACCGCAGTGAAGGTTCCATTCACCAAAATTAATGTTCTCT	803255		
666	CGGACAGCTTATATATGCAAAAGTGTGATTTTCACTCAATCCGTTGACCTCAACCT	725		
803256	AGGCTATATTCATTAACGCAAAATGTGACGTTCCCGAGAAAGCAAAATCATACAT	803315		

Db 1729 AATATGGCTTCGTCGCTTGGCAGCGTGAACCCGCAAGGCGGCTTCTGCTGCA 1788
 QY 1051 TTAGCCCTCACCACAAATAGTAGAGAGGCGATCCAAAGATCTTCTCCACTATTAA 1110
 Db 1789 CTGGCTGTAGAGCAAAACCAAGATCCGACAGATCCAGCATCTTCAATAGTACTAA 1848

RESULT 5 PCT-US92-10421-1

; Sequence 1, Application PC/TUS9210421

GENERAL INFORMATION:

APPLICANT: Robert, Joseph
 APPLICANT: MacCallister, Thomas W
 APPLICANT: Sethuraman, Natarajan
 APPLICANT: Freeman, Abbie G
 TITLE OF INVENTION: GENETICALLY ENGINEERED GLUTAMINASE AND
 TITLE OF INVENTION: ITS USE IN ANTI-VIRAL AND ANTICANCER THERAPY
 NUMBER OF SEQUENCES: 10
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Banner, Birch, McKie and Beckett
 STREET: 1001 G Street N.W.
 CITY: Washington, D.C.
 COUNTRY: U.S.A.
 ZIP: 20001

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: PCT/US92/10421
 FILING DATE: 19921204

CLASSIFICATION:

PRIOR APPLICATION DATA:
 APPLICATION NUMBER: DE P4140003.8
 FILING DATE: 04-DEC-1991

ATTORNEY/AGENT INFORMATION:

NAME: Kagan, Sarah A

REGISTRATION NUMBER: 32,141
 REFERENCE/DOCKET NUMBER: 00100.41200
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 202-508-9100

TELEFAX: 202-508-9299

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 1014 base pairs

TYPE: NUCLEIC ACID

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

HYPOHETICAL: NO

ANTI-SENSE: NO

ORIGINAL SOURCE:

ORGANISM: PSEUDOMONAS 7A

PCT-US92-10421-1

Query Match 19.0%; Score 215.8; DB 5; Length 1014;

Best Local Similarity 52.6%; Pred. No. 1.9e-61;

Matches 519; Conservative 0; Mismatches 462; Indels 6; Gaps 2;

QY 128 CCAAGTACTATCTAGCCACAGAGGACCATCGCTGGTTCGGGGGAATCTAGCGTCA 187
 Db 29 CCAAGTACTATCTAGCCACAGAGGACCATCGCTGGTTCGGGGGAATCTAGCGTCA 88
 QY 188 AGAG---TAGCTACTGCTGAGAGCATCACCCTGTGATAGCTTTCTTGCAAGCCGTCCTG 244
 Db 89 ACAGCGCCACCTACCAAGCTGCGCTGCGACAGCTGATTCGCGGCTGCGG 148
 QY 245 CCATCAAGAGCTTCAAGGATCAAGATCTCAAGCATTCGCTCCCAAGAGA 304
 Db 149 AGCTGGCCGACCTGCGCAATCTGCGCGGCGAGAGTGTGATGCAATGCGCTCGCAAGACA 208
 QY 305 TGACGGGTAAAGTGTGCTTAAGTACCAAGCGTGTCAATGAGCTCTCTGCGCCCAAAAG 364

Db 209 TCACCAACGACGACCTGCTCAAGCTGGCAAGAGCGGTGGCGAGCTGGCCGACAGCAATG 268
 QY 365 AGACCGAAGCCGTGATCATACCATGGAATCAACATGGAAGAGACCGCTTCTTCC 424
 Db 269 ACCTGATGTCATCTCATCCATCCATGGAACCGACACCTGGGAAGAACCGCTTCTTCT 328
 QY 425 TCACCTCAGCGTGAAGAAAGCCAAACCTGTGCTTGTAGAGCGCATGCTCCAGCT 484
 Db 329 TGAACCTCTGTGGAAGAAAGACCGACAGCGCATCTGCTGTGCTTGTCAATGCGCGGGA 388
 QY 485 CTTCATAGTGTGATGAGCCCATGATCTATTAACGCGTGAATGTATGCGATCAACA 544
 Db 389 CCGCATGTCCGCCGACGATGCTCAACCTGTACAAAGCGCTGTGCGCGACGACGAACA 448
 QY 545 AAGCTCTACTAACAAGAGAGTGTGATGATGAAGATGATGATGATGATGATGATGATG 604
 Db 449 AGGACTGCGCGGCAAGAGGCTGTGCTGTGACATGAAAGAGAGATGATGATGATGATG 508
 QY 605 AAGCAGCAAGCTCAACACCAACCGCATCAATGATTTGCTTGGCCCAACAGGTAATA 664
 Db 509 ACCTGAGCAATGATCAATCAATCAAGCCGAAGCTTC---AAGAGCGCTGGGGCGCG 565
 QY 665 TCGCAGAGTCTATTATGCAAGTCAATTTCACTCAATCGCTTGCACCTCACACC 724
 Db 566 TGGCATGTGTGTGAAGGCAAGTCTGATCTGCTGCTGCGCGCCAGCGCCACGCG 625
 QY 725 TTGCAAGTGTGATATTAGCAAAATCGAAGACTCCCGAGATGATTTCTTACG 784
 Db 626 TCAACTCGAGTGTGATCAATCAAGAGATCAAGATCAAGATCAAGATCAAGATCAAGAT 685
 QY 785 CTCACCCGATGATGATGATTTTATGCAATGCAAGCTTCAGGCAAGAGCCAAAGGA 844
 Db 686 GCTATGCAAGCTCAACGACGCGCTTACAGCGCTTGGCACAGAGCGCGCCAAAGCGC 745
 QY 845 TCATCATGCAAGGATGAGGCAATGGAACCTTTCCCTTACCAAAATGCTTTGAAA 904
 Db 746 TGATCATGCGCGGACCGGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 805
 QY 905 AAGCAGCAATCAAGGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 964
 Db 806 AGCTGCGCAAGAGCGCGCTGCAATCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 865
 QY 965 CCAAGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1024
 Db 866 TCGCTAGCGCGAGAGCGCGCGCGAGCAAGAGCAAGAGCAAGAGCAAGAGCAAGAG 925
 QY 1025 CTCAAAAGCCGAGAGTGTCTTATGTTAGCCCTCACCAAACTAGTATGAGAGGCGA 1084
 Db 926 CGCAGAGGCGCGCATCTCTGCGGATGCTGCGCAATGACCAAGAGCCAGAGCAAGAG 985
 QY 1085 TCCAAAAGATCTTCTCAACCTATTAT 1111
 Db 986 TGCAGCGCATTTCTGCGGAATCTGAT 1012

RESULT 6

US-08-050-482A-1

; Sequence 1, Application US/08050482A

; Patent No. 6312939

GENERAL INFORMATION:

APPLICANT: ROBERTS, Joseph
 APPLICANT: MACALLISTER, Thomas W.
 APPLICANT: SETHURAMAN, Natarajan
 APPLICANT: FREEMAN, Abbie G.
 TITLE OF INVENTION: GENETICALLY ENGINEERED GLUTAMINASE AND
 TITLE OF INVENTION: ITS USE IN ANTI-VIRAL AND ANTICANCER THERAPY
 NUMBER OF SEQUENCES: 22
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: FOLEY & LARDNER
 STREET: 3000 K Street, N.W.
 CITY: Washington
 STATE: D.C.


```

COUNTRY: U.S.A.
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/050,482A
FILING DATE: 25-Apr-1995
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US92/10421
FILING DATE: 04-DEC-1992
ATTORNEY/AGENT INFORMATION:
NAME: Bent, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 023032/0106
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 672-5300
TELEFAX: (202) 672-5399
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1017 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1011
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-08-050-482A-1

Query Match      18.6%; Score 210.8; DB 4; Length 1017;
Best Local Similarity 52.9%; Pred. No. 8.8e-60;
Matches 524; Conservative 0; Mismatches 457; Indels 9; Gaps 3;

QY 128 CCCAAGGACTATTCCTAGCCACAGAGGACACCATCGCTGTCGGGGGAATCTAGCGTCA 187
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 29 CCAAGGTGTGATCTCTCCGACCGCGGCACCATCGCGGGCGTGGCCAGCGCGCCCA 88
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 188 AGAG---TAGCTACTGTCTGTGAGCAGTCAACCGTTGATTAAGCTTCTTGACGCGCTCCCTG 244
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 89 ACAGCGGCACCTACACAGCGTCCAGAGTGTGGCGTCGACAAAGCTGATTCGCGCGTGGCGG 148
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 245 CCATCAACGACCTTAGCCACCATCAAGGGTGAACAGATCTCAAGCATTTGGTCCCAAGAGA 304
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 149 AGCTGGCGGACCTGGCCCAATGTGCGCGGAGCAGGTGATGACAGATCGCTCCGAAAGCA 208
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 305 TGAGGGGTAAAGTGTGGCTTAAATAGCAAGGGTGTCAATGAGCTTCCTCGCCCAAAAG 364
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 209 TCACCAAGCAGACCTGTCTAAGCTGGGCAAGCGCGTGGCGGAGCTGGCCGACGACATG 268
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 365 AGACCGAAGCGGTGATCATACCATGGAAGTGAACACATGGAAGAGACCGCTTTCTTCC 424
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 269 ACCTGATGTCATCATCATACCATGACGACGACACCTGGAAGAAACCGCTACTTTT 328
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 425 TCACCTCAAGCGGGAAGCAAAACCTGTCTGCTTGTAGAGCGCATGCTCGAGGCT 484
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 329 TGACCTCGGTGAAAAGACCGACAGCGCATCTGCTGTGGTTCATGCGCCCGGCA 388
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 485 CTTCGATGAGTGTGATGGCCCATGAAATCTTAAATAGCGCTGAAATGTAGCATCAACA 544
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 389 CGGCATGTGCGCGACGCGCATCTCAACCTGTACAAAGCGCGTGGCGCGACGACACA 448
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 545 AAGCCCTACTAACAAGAGAGTGTGATGTGTAAGCATGAGATTACAGCGCCGACAGAG 604
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 449 AGGACTCGCGGCGGAGGCGTGTGTGTGATGACATGACGAGATCCAGATCCGCGCGTG 508
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 605 AAGCGCAAGGTCTCAACACCGGACGTCATATGCAATTTGGTTGGCCCAACAGGTAANA 664
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 509 AGGTAGCAAGTGTGATCAACATCAAGACCGAAGCCTTC---AAGAGCGGCTGGGCGCCG 565

```

```

QY 665 TCGGCACTATCTATTATGCGAAAGTCAGTATTTCATCATCCGTTCCACCTCACACC 724
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 566 TGGGATGTGTGTGGAAGGCAAGTCGTAATGTTGCTGCGCTGCGCGGCAAGCCGACACG 625
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 725 TTGCAAGTGAAGTTTATATTACCAAAATGGAAGAACTCCCGAGATGCAATTTCTTTAG 784
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 626 TCAACTCCGATTCGACATCAAGCAAGATCAGAGCTGCGCCAGGTGACATCGCTTACA 685
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 785 CTACCCCGATGATCTACTGTTTATGTCATGACAGCCCTTAGAGCAGGAGCCAGAGAA 844
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 686 GCTATGGCAAGTCAACCGACAGCGCTTACAAAGCCCTTGACACAGAGCGCGGAGGCGC 745
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 845 TCATCTATGACAGGCTAGGCAATGGGAACCTTTCCCTTTGACTCAAAATGCTTTGAA 904
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 746 TGATCATGCGCGGACCGGCATGCTGCTGCTGCTGCGGGTGTGACGACCTGACAG 805
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 905 AAGCAGCAAAATCAGAGCTAGTCTGCTGCGAAGCTTAGAGTGGC---AGTGGTTCA 961
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 806 AGCTGCGCAAGAGCGCGCTGACAGATCATTCCTGCTCCACGTCACACAGGCGGCTTGC 865
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 962 CCACCCAGAGGCTGATGATGATTAAGAACTTGTGTTGTGCTACAGAGACTCA 1021
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 866 TGCTGCTAGCGCGGACGCGCCGACGACAAAGACGACTGGGTGCTGCGCCACGACCTGA 925
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1022 ACCCTCAAAAGCCAGAGTGTCTTATGTAAGCCCTACCCAAACTAGTATAGAGG 1081
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 926 ACCGCGAAGAGCGCCGCTCTGCGGATGTGGCAATGACCAAGACCCAGGACAGCAAG 985
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1082 CGATCAAAAGATCTTCTCCACCTAATTAAT 1111
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 986 AGCTGACGCGCATTTTCTGGGAATACTGAT 1015
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 7
US-09-252-991A-5793
Sequence 5793, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
PRIOR FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 3142
SEQ ID NO 5793
LENGTH: 1182
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-5793

Query Match      17.7%; Score 200.4; DB 4; Length 1182;
Best Local Similarity 52.5%; Pred. No. 2.8e-56;
Matches 513; Conservative 0; Mismatches 456; Indels 9; Gaps 3;

QY 128 CCCAAGGACTATTCCTAGCCACAGAGGACACCATCGCTGTCGGGGGAATCTTAGCGTCA 187
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 197 CCAAGGTGTGATCTCTCCGACCGCGGCACCATCGCGGGCGCGGCGGCGCGCA 256
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 188 AGAG---TAGCTACTGTCTGTGAGCAGTCAACCGTTGATTAAGCTTCTTGACGCGCTCCG 244
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 257 ACAGCCGACCTTACACCGCGGCAAGGTGCGCGTGTGACAGTGTGCGCAGAGTGC 316
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 245 CCATCAACGACCTTAGCCACCATCAAGGGTGAACAGATCTCAAGCATTTGGTCCCAAGAGA 304
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 317 AGCTCAAGGACATGCGCAACGATCGCGGCGAAGAGTGTTCAGATCGCTCGGAAGCT 376
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 305 TGAGGGGTAAAGTGTGGCTTAAATAGCAAGGGTGTCAATGAGCTTCCTGCGCCAAAG 364

```


QY 1082 CGATCCAAAGATCTTCT 1099
 Db 254 AGCTGCAAGGATCTTCT 237

RESULT 9

US-09-328-352-2336
 ; Sequence 2336, Application US/09328352
 ; Patent No. 6562958
 ; GENERAL INFORMATION:
 ; APPLICANT: Gary L. Breton et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
 ; FILE REFERENCE: GTC99-03PA
 ; CURRENT APPLICATION NUMBER: US/09/328,352
 ; CURRENT FILING DATE: 1999-06-04
 ; NUMBER OF SEQ ID NOS: 8252
 ; SEQ ID NO 2336
 ; LENGTH: 1125
 ; TYPE: DNA
 ; ORGANISM: Acinetobacter baumannii
 US-09-328-352-2336

Query Match 14.1%; Score 159.2; DB 4; Length 1125;
 Best Local Similarity 50.3%; Pred. No. 14e-42;

Matches 504; Conservative 0; Mismatches 483; Indels 15; Gaps 4;

QY 121 GCTAAACCCCAAGTACTATCTTACGACAGAGGACACATCGTGTTCGGGGGAATCT 180
 Db 127 GCAAAAACAAATGTTGTTGTTGCTGCTACGCGGTACATTCGCGGTGGGCAAGC 186
 QY 181 AGCGTCAGAGT---AGCTACTCTGCTGGAGCAGTACCGCTTGATTAAGCTTCTGACGC 237
 Db 187 TCAGAAATATGCACTTATATACGCGGCAAAAGTTCAGTTGATCTTATCAATGCA 246
 QY 238 GTCCCTGCCATCAACGACCTACACCATCAAGGGTGAAGATCTCAAGATTTGGCTCC 297
 Db 247 GTTCTCTCAATTCAGATTTGGCGATATCTGTATCCAGCAATTCAGATGACTTCT 306
 QY 298 CAAAGATGACGGGTGAGGTGCTTAACTAAGCCAGCTGTCATAGAGCTCCGCTGCC 357
 Db 307 GAAAGTATATCTGACAGGAATTTATCAAAATGCTGCTCAAGTAAATGACTTGTAAA 366
 QY 358 CAAAAGAGACGGAAGCGGTATCATACCATGAGTGAACCATGAGAGAGAGCCGCT 417
 Db 367 AAGCACTGTAAATGGCTGTGTATACACGCTACAGATCTTGAAGAAACAGCA 426
 QY 418 TTCTTCTCAACCTACAGGTGAAAAAGCCAAACCTGTCTGCTGTAGGCGCATGGCT 477
 Db 427 TTTTCTTAATCTGTGTATCATATGATTAACCAATCGTACTGTTGGCTCAATGGC 486
 QY 478 CCAGGCTTTCAGAGTGTGATGGCCCATGAATCTCTTAACGCGGTGAATGAGG 537
 Db 487 CCAATCACTGCTTTTACAGAGATGTCTCACTTAACCTTATAGGCTGTGCAATGGCC 546
 QY 538 ATCAACAAGCGCTTACTAACAAGAAGTGGTATGATGAAGATGAGATGACAGC 597
 Db 547 GCTTCTGATGATGAAAAAATAAAGGCTTATGGTTCTCATGAAGACTCTAATTTTGGCT 606
 QY 598 GCCAGAGAGGACCAAGCTCAACACCAAGCTCAATGATTTGCTTCCGCCAACACA 657
 Db 607 GCTGTGATGATTAAGGATTAACATTCATACAAATGCTTTTGAAGCC---AATGG 663
 QY 658 GGTAAATCGGACAGTCTATATGCAAGATGCAATTTCACTCAATCCGTTGACCT 717
 Db 664 GGTGCTTTGGTACACTTGTGAAGCAAAACCATTTGGTTTGAACATCTGTTAAAGCT 723
 QY 718 CACACCTTGCAGAGTGTGATAT-----TAGCAAAATCGAAGAACCTCCAGAGTC 771
 Db 724 CATACCAATGCTTCAAGATTTAATATGAAATAATTAAGGTGAGTCTTCCACAGTA 783
 QY 772 GATATCTTACAGCTCAACCCGATGATGATGATGTTTATGCAATGACGCCCTCAGGCA 831

Db 784 CAAATGCTTTATGTTGATCTATGCTTCTGATGCTTATGAAGCATATGAAGAACT 843
 QY 832 GGAGCCAAAGATATCATCTGACGATGGCAATGGAA---CCCTTCCCTTGACT 888
 Db 844 GGCGTAAAGCATTAATTTATCATGAGTACAGGTAAAGGTTCTGTAAGCAAAATATTCGT 903
 QY 889 CAAATGCTCTGAAAAAAGCAGCAATACAGCGTATGCTGCGTGAAGCTCTGAGAG 948
 Db 904 CCAACATTCAAACCTTCATACCAAAAAGGATATTCAAATATATCGCTCATCGCGCTT 963
 QY 949 GCGAGTGTTCACACCAACCAAGAGGCTGAATGATATGAAGAACTTGGTTTGCT 1008
 Db 964 CCACAGGTTTGTATTTACGTATGACAGAACTGACTTAATATGTTGGTAGCA 1023
 QY 1009 ACAGAGCTTCACACCTCAAAAGCAAGAGTCTTCTATATGTTAGCCCTCAGCAAACT 1068
 Db 1024 GCTATGATTTAAATCTCAAAAGGCGCTCTTCTGCGGCTGTGGCACTTACCAAAAC 1083
 QY 1069 AGTATAGAGAGGCGATCCAAAAGATCTTCTCACCCTATTAA 1110
 Db 1084 AACGATGCGAAGAAATTCACAGCATGTTCTGCGAGTACTAA 1125

RESULT 10

US-09-252-991A-5773
 ; Sequence 5773, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252,991A
 ; CURRENT FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 5773
 ; LENGTH: 2748
 ; TYPE: DNA
 ; ORGANISM: Pseudomonas aeruginosa
 US-09-252-991A-5773

Query Match 13.0%; Score 147.4; DB 4; Length 2748;
 Best Local Similarity 51.8%; Pred. No. 2e-38;

Matches 384; Conservative 0; Mismatches 351; Indels 6; Gaps 2;

QY 362 AAGAGCCGAAGCGGTATCATACCCATGGAATGACACATGGAAGAGACCGCTTCT 421
 Db 3 AAGAGCGTACGCGATCTGATACCCAGGCTACCGACACCTGGAAGAGACCGCTTCT 62
 QY 422 TCCATCAACCTCAAGGTAAGAAAGCCAAACCTGTGCTTGTATGCGCCCATGGCTCAG 481
 Db 63 TCTGACCTCTGTGAGACACACCGAAGCCATATGCTGTGTGCTGATGCGCCGG 122
 QY 482 GCTTTCATGATGCTGATGATGCCCCCATGAATCTCTATACGCGGTGAATGAGATCA 541
 Db 123 GCACCGCATGTCCGCGCAGCGAGATGCTCAACCTGTAACAGCGGTGCGGTGGCG 182
 QY 542 ACAAGCTCTACTTAACAAGAGATGCTGATTTGATGATGAAGATGAGATTCAGCGCCCA 601
 Db 183 ACAAGTGGGACGCGGCGAAGGCGCTGTGATCACCATGAAAGAGAGATCTCTCCGCC 242
 QY 602 GAGAGGACCAACAGCTCAACACACCGAGTCAATGATTTGCTTCCGCCAACACAGTA 661
 Db 243 GCGACGGAGCAAGATGCTCAACATCAAGACGAGGTT---CAAGAGCCGTTGGGCC 299
 QY 662 AATGCGACAGCTATATATGCAAGATGCAATGATTTCACTCAATCCGCTTGCACCTGCA 721
 Db 300 CGCTGGGATGCTGTGAGAGGCAAGAGCTACTGTTCCGCGCACCGGTGAAGCGGACCA 359

; LOCATION: (1)...(363)
US-09-613-303-11

Query Match 6.4%; Score 72.6; DB 4; Length 366;
Best Local Similarity 94.9%; Pred. No. 4.5e-14;

Matches 75; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY	1	ATGGGACAGCAGCCATCATCATCATATGACAGCGGCTGTGCGCGCGGACGCCAT	60
Db	1	ATGGGACAGCAGCCATCATCATCATATGACAGCGGCTGTGCGCGCGGACGCCAT	60
OY	61	ATGGCTAGCATGACTGTG	79
Db	61	ATGGCTAGCATGACTGTG	79

Search completed: September 4, 2003, 18:15:29
Job time : 102 secs


```
NAME/KEY: misc_feature
```

```
NAME/KEY: misc_feature
```

OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (44416)..(44416)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (44905)..(44905)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (44975)..(44975)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (45593)..(45593)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (45732)..(45732)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (47036)..(47036)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51334)..(51334)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51602)..(51602)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51786)..(51786)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51805)..(51805)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (55369)..(55369)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (65309)..(65309)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (65313)..(65313)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (80024)..(80024)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (100091)..(100091)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (102696)..(102696)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (105121)..(105121)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (107248)..(107248)
OTHER INFORMATION: n equals a, t, g or c

FEATURE:
NAME/KEY: misc_feature
LOCATION: (117136)..(117136)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (119750)..(119750)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (119924)..(119924)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (120038)..(120038)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (121344)..(121344)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (122167)..(122167)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (122336)..(122336)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (131340)..(131340)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (131360)..(131360)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (139910)..(139910)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (140398)..(140398)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (142750)..(142750)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145058)..(145058)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145171)..(145171)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145942)..(145942)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (147197)..(147197)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (150841)..(150841)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (152500)..(152500)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:


```

Db      697 AGGAGCGCTGATATATGCAAAAGCGCGCTATTACATGAGCGCTTGAAAAACACACAC 756
      726 TGCAGTGTGATTTGATTTAGCAAAATGAA---GACTCCCGAGAGTCGATTTCTTTA 782
      757 AGAGGCGCAATTTTCCCTTTTACACACTCAAAACCCCTGCTTAAGTGATTTATTTA 816
      783 GCCTCACCCCGGATGATGATGATTTAGTCATGACAGCCCTTGAGCAGAGACCCAAAG 842
      817 CACGATGCTGGCAGACCCCTGATTTATTCACAGAGCCTAACTGATGCAAAAG 876
      843 AATCATTCATGAGCAGATGAGGCAATGAGACCCCTTTGACTCAAAATGCTCTGA 902
      877 CATTGTGATGAGCGGGGTGGGTAATGGAATGAGCGCTGGTTTAAAGCATCA 936
      903 AAGAGCGCAATGAGCGGTGATGCTGCTGCAAGCTCTAGAGTGGCAGTGTTCAC 962
      937 AGAAGCGAGCCAAATGGGGGTGTTATTTGTTCTGACAGGTAATAGCGGTGAGAT 996
      963 CACCCAGAGGCTGAAGTGATGATPAGAACTGTTTGTGCTACAGAGTCTCA 1022
      997 TA-----CTTACAGCGGATGATGATGACAGCGCTTCATCAGAGGCAATTTAA 1047
      1023 CCTCAAAAGCGAGGCTCTTATTTAGTACCCCTCACCAGCTAGTATAGAGAGC 1082
      1048 CCCCCCAAGAGTGAAGGCTCTTTACACTCGCTTTACTAAACAAATATRAAGAAA 1107
      1083 GATCAAAAGATCTTCTCCACTATTAA 1110
      1108 AATCCAGAAATGTTGAAGATTTGA 1135

```

RESULT 4

```

US-09-842-628-1
: Sequence 1, Application US/09842628
: Patent No. US20020064862A1
: GENERAL INFORMATION:
: APPLICANT: ROBERTS, JOSEPH
: APPLICANT: MACALISTER, THOMAS W.
: APPLICANT: SETHURAMAN, NATARAJAN
: APPLICANT: FREEMAN, ABIE G.
: TITLE OF INVENTION: GENETICALLY ENGINEERED GLUTAMINASE AND ITS USE IN
: TITLE OF INVENTION: ANTIVIRAL AND ANTICANCER THERAPY
: FILE REFERENCE: 023032/0108
: CURRENT APPLICATION NUMBER: US/09/842,628
: CURRENT FILING DATE: 2001-04-27
: PRIOR APPLICATION NUMBER: 08/050,482
: PRIOR FILING DATE: 1995-04-25
: PRIOR APPLICATION NUMBER: PCT/US92/10421
: PRIOR FILING DATE: 1992-12-04
: PRIOR APPLICATION NUMBER: DE P 4140003.8
: NUMBER OF SEQ ID NOS: 22
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 1
: LENGTH: 1014
: TYPE: DNA
: ORGANISM: Pseudomonas sp.
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (1)..(1008)
US-09-842-628-1

```

```

Query Match      19.0%; Score 215.8; DB 9; Length 1014;
Best Local Similarity 52.6%; Pred. No. 2.9e-63;
Matches 519; Conservative 0; Mismatches 467; Indels 6; Gaps 2;

```

128 CCGAAGTACTATCTTACACAGAGGACCACTGCTGTGCGGGGAATCTAGCTCA 187
 29 CCAAGCTGATCTGCGACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCA 88
 188 AGAG---TAGCTACTCTGCTGAGAGAGTACCGGTGATAGAGTCTTTCAGCGCGCTG 244

```

Db      89 ACAGCGCACCTACAGGCTGCCAAGGTGGCGTCGACAAAGCTGATTTGCCGGCTGCCG 148
      245 CCATCAACGACCTTAGCCACCATCAGGCTGAACAGATTCACATTTGGCTCCCAAGAGA 304
      149 AGCTGGCGGACCTGGCCAAATGTGCGGCGGAGGATGATGAGATCGCTCCGAAAGGA 208
      305 TGAAGGTAAGTGTGGCTTAACTAGCCCAAGCGTGTCAATGAGCTCTGCGCAAAAG 364
      209 TCACCAAGACGACCTCTCAAGCTGCGCAAGCGGCGGCGGCGGCGGCGGCGGCGGCA 268
      365 AGACCGAAGCGGTGATCATCACCATGGAAGTGAACCTGAGAGAGAGACCGCTTTTCC 424
      269 ACGTATGAGCATGATCATCACCATGGAAGTGAACCTGAGAGAGAGAGAGAGAGAGAG 328
      425 TCACCTCAGCGTGAAGAAAGCAAAAGCTGCTGCTGAGGCGCATGCTGCAAGCT 484
      329 TGAACCTGCTGGAAGAAAGCAAAAGCTGCTGCTGAGGCGCATGCTGCAAGCT 388
      485 CTTCATGAGTGTGATGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCA 544
      389 CCGCATGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCA 448
      545 AAGCTCTACTAACAAGAGAGTGTGATGATGAGAGAGAGAGAGAGAGAGAGAGAGAG 604
      449 AGGACTCCGCGGCGGAGGCGGCTGCTGAGACCATGAGAGAGAGAGAGAGAGAGAGAG 508
      605 AAGCGACCAAGCTCAACACACCGCATGATGATGATGATGATGATGATGATGATGATG 664
      509 ACGTAGAGAGTGCATCAACATCAAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 565
      665 TCGGCAAGTCTATTATGCAAAAGTGAATTTTCAATTTTCAATTTTCAATTTTCAATTT 724
      566 TGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 625
      725 TTGCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 784
      626 TCACTCGAGTGCATCAACATCAAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 685
      785 CTCACCGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 844
      686 GGTATGCGCAAGCTCAACACCGCATGATGATGATGATGATGATGATGATGATGATGAT 745
      845 TCATCATGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 904
      746 TGAATCATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 805
      905 AAGCAGCCAAATCAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 964
      806 AGCTGCGCAAGAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 865
      965 CCCAAGAGGCTGAAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1024
      866 TGGCTAAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 925
      1025 CTCAAAAGCGAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1084
      926 CGCAGAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 985
      1085 TCCAAAAGATCTTCTCCACTATTAA 1111
      986 TGCAGCGCATTTCTGGAATGATGAT 1012

```

RESULT 5

```

US-09-876-348A-34
: Sequence 34, Application US/09876348A
: Patent No. US20020172951A1
: GENERAL INFORMATION:
: APPLICANT: Horvath, K. L. and Myers, K. L.
: TITLE OF INVENTION: Nucleic Acid Sequences Encoding Type III Tenebrio
: TITLE OF INVENTION: Antifreeze Proteins and Method for Assaying Activity.
: FILE REFERENCE: RB-125-RI
: CURRENT APPLICATION NUMBER: US/09/876,348A

```

Query Match	9.0%	Score 102;	DB 10;	Length 543;
Best Local Similarity	88.1%;	Pred. No. 3.8e-24;		
Matches 111;	Conservative 0;	Mismatches 15;	Indels 0;	Gaps 0;
QY	1	ATGGGACAGACCCATCATCATCATCATCATCAGCAGCGCGCTGGTGGCGGCGGACGCAT	60	
Ddb	64	ATGGGACAGACCCCATCATCATCATCATCATCAGCAGCGCGCTGGTGGCGGCGGACGCAT	123	
QY	61	ATGGCTTGACATGACTGGTGGACAGCAAAATGGTTCGGGATTCAGCGCTCTTTTGTATG	120	
Ddb	124	ATGGCTTGACATGACTGGTGGACAGCAAAATGGTTCGGGATTCAGCGCTCTTTTGTATG	183	
QY	121	GCTAAA	126	
Ddb	184	CAGAAA	189	

QY 61 ATGGCTAGCATGCTGTGGACAGCAATGGTGGGATCCAGCCCTCTGTTTGATG 120
|||||
Db 124 ATGGCTAGCATGCTGTGGACAGCAATGGTGGGATCCCTCACCAGCAAGATTA 183
QY 121 GCTAAA 126
|||
Db 184 CAGAAA 189

RESULT 13

US-09-876-796A-30
; Sequence 30, Application US/09876796A
; Patent No. US20020173024A1
; GENERAL INFORMATION:
; APPLICANT: Horwath, K. L. and Easton, C. M.
; TITLE OF INVENTION: Nucleic Acid Sequences Encoding Type III Tenebrio
; FILE REFERENCE: RB-125-SEQ
; CURRENT APPLICATION NUMBER: US/09/876,796A
; PRIOR FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: 60/210,446
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: Microsoft Word
; SEQ ID NO 30
; LENGTH: 543
; TYPE: DNA
; ORGANISM: Tenebrio molitor
; OTHER INFORMATION: His-tagged, Signal minus, Clone 3.4
US-09-876-796A-30

Query Match
Best Local Similarity 9.0%; Score 102; DB 10; Length 543;
Matches 111; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 1 ATGGGCAGCAGCATCATCATCATCATCATAGCAGCGGCTGTGGCGGCGGAGCCAT 60
|||||
Db 64 ATGGGCAGCAGCATCATCATCATCATCATAGCAGCGGCTGTGGCGGCGGAGCCAT 123
QY 61 ATGGCTAGCATGCTGTGGACAGCAATGGTGGGATCCAGCCCTCTGTTTGATG 120
|||||
Db 124 ATGGCTAGCATGCTGTGGACAGCAATGGTGGGATCCCTCACCAGCAAGATTA 183
QY 121 GCTAAA 126
|||
Db 184 CAGAAA 189

RESULT 14

US-09-876-796A-38
; Sequence 38, Application US/09876796A
; Patent No. US20020173024A1
; GENERAL INFORMATION:
; APPLICANT: Horwath, K. L. and Easton, C. M.
; TITLE OF INVENTION: Nucleic Acid Sequences Encoding Type III Tenebrio
; FILE REFERENCE: RB-125-SEQ
; CURRENT APPLICATION NUMBER: US/09/876,796A
; PRIOR FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: 60/210,446
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: Microsoft Word
; SEQ ID NO 38
; LENGTH: 543
; TYPE: DNA
; ORGANISM: Tenebrio molitor
; OTHER INFORMATION: His-tagged, Signal minus, Clone 7.5
US-09-876-796A-38

Query Match
Best Local Similarity 9.0%; Score 102; DB 10; Length 543;
Matches 88.1%; Pred. No. 3.8e-24;

Matches 111; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 1 ATGGGCAGCAGCATCATCATCATCATATAGCAGCGGCTGTGGCGGCGGAGCCAT 60
|||||
Db 64 ATGGGCAGCAGCATCATCATCATCATATAGCAGCGGCTGTGGCGGCGGAGCCAT 123
QY 61 ATGGCTAGCATGCTGTGGACAGCAATGGTGGGATCCAGCCCTCTGTTTGATG 120
|||||
Db 124 ATGGCTAGCATGCTGTGGACAGCAATGGTGGGATCCCTCACCAGCAAGATTA 183
QY 121 GCTAAA 126
|||
Db 184 CAGAAA 189

RESULT 15

US-10-228-063-43
; Sequence 43, Application US/10228063
; Publication No. US20030135885A1
; GENERAL INFORMATION:
; APPLICANT: Lananhan, Mike
; TITLE OF INVENTION: Self-Processing Plants and Plant Parts
; FILE REFERENCE: 109846.317
; CURRENT APPLICATION NUMBER: US/10/228,063
; CURRENT FILING DATE: 2002-12-12
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 43
; LENGTH: 1436
; TYPE: DNA
; ORGANISM: Thermotoga neapolitana
US-10-228-063-43

Query Match
Best Local Similarity 9.0%; Score 101.6; DB 12; Length 1436;
Matches 125; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY 1 ATGGGCAGCAGCATCATCATCATCATATAGCAGCGGCTGTGGCGGCGGAGCCAT 60
|||||
Db 1 ATGGGCAGCAGCATCATCATCATCATATAGCAGCGGCTGTGGCGGCGGAGCCAT 60
QY 61 ATGGCTAGCATGCTGTGGACAGCAATGGTGGGATCCAGCCCTCTGTTTGATG 120
|||||
Db 61 ATGGCTAGCATGCTGTGGACAGCAATGGTGGGATCCCTCACCAGCAAGATTA 183
QY 121 GCTAAACCCCAAGTGTCTTCTAGCCAGAGGACACCATGCC 164
|||||
Db 121 GAGATCCCAAGTGTCTTCTAGCCAGAGGACACCATGCC 164

Search completed: September 4, 2003, 19:20:51
Job time : 352 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 4, 2003, 16:12:46; Search time 2338 Seconds
(without alignments)
11778.003 Million cell updates/sec

Title: US-09-937-982-3

Perfect score: 1133
Sequence: 1 atggcagcagcattcatca.....aagaagggaattctttac 1133

Scoring table: IDENTITY-NUC
Gapop 10.0, Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database: EST:
1: em_estpa:*
2: em_esthm:*
3: em_estln:*
4: em_estmv:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hnc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hnc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_esthm:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rtd:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	155.2	13.7	983	29	BZ568327 pacs2-164
2	136	12.0	724	28	AC0990947 RfC01794
3	116.8	10.3	766	28	AC0500605 V36C6 mtn
4	110.8	9.8	757	28	BH376509 AG-ND-137

Result	Score	Query Match	Length	DB ID	Description
1	155.2	13.7	983	29	BZ568327 pacs2-164
2	136	12.0	724	28	AC0990947 RfC01794
3	116.8	10.3	766	28	AC0500605 V36C6 mtn
4	110.8	9.8	757	28	BH376509 AG-ND-137

ALIGNMENTS

RESULT 1
BZ568327/c
LOCUS pacs2-164_7476.y2 pacs2-164 Pseudomonas aeruginosa genomic clone
DEFINITION
ACCESSION BZ568327.1 GI:27201144
VERSION
KEYWORDS
SOURCE
ORGANISM
Pseudomonas aeruginosa
Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
Pseudomonadaceae; Pseudomonas.
1 (bases 1 to 983)
Spencer, D.H., Raymond, C.K., Smith, E.E., Sims, E.E., Hastings, M.,
Burns, J.L., Kaul, R. and Olsen, M.V.
Whole-Genome-Sequence Variation among multiple isolates of
Pseudomonas aeruginosa library
J. Bacteriol. (2002) In press
Contact: Chris K. Raymond
Genome Center
University of Washington
Box 352145, Seattle, WA 98105-2145, USA
Tel: 2062216954
Fax: 2066857244
Email: ckraymond@u.washington.edu
Class: Shotgun
Location/Qualifiers

FEATURES

Query Match	13.7%	Score 155.2;	DB 29;	Length 983;
Best Local Similarity	58.7%	Pred. No. 1.2e-34;		
Matches 304; Conservative	0;	Mismatches 210;	Indels 4;	Gaps 2.

187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000

RESULT 2	724 bp	DNA	linear	GSS 14-AUG-2000
LOCUS	AO990947/c			
DEFINITION	AO990947	724 bp	DNA	linear
	RFc01794	Photorhabdus luminescens strain W14 M13 library		
	Photorhabdus luminescens genomic clone PLC01794, genomic survey			
	sequence.			
ACCESSION	AO990947			
VERSION	AO990947.1	GI:9649541		
KEYWORDS	GSS.			
SOURCE	Photorhabdus luminescens			
ORGANISM	Photorhabdus luminescens			
	Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;			
	Enterobacteriaceae; Photorhabdus.			
REFERENCE	1. (bases 1 to 724)			
AUTHORS	French-Constant, R.H., Waterfield, N., Burland, V., Perna, N.T.,			
	Daborn, P.J., Bowen, D., and Blattner, F.R.			
TITLE	A genomic sample sequence of the entomopathogenic bacterium			
	Photorhabdus luminescens W14: potential implications for virulence			
JOURNAL	Appl. Environ. Microbiol.	66 (8),	3310-3329	(2000)
MEDLINE	20378633			

10919786

CONTACT: French-Constant RH
Department of Biology and Biochemistry
University of Bath
South Building, Bath BA2 7AY, UK
Tel: (44) 1225 826621
Tel: (44) 1225 826729
Fax: (44) 1225 826779
Email: bsr1fc@bath.ac.uk
This is one of 2,122 random reads from the M13 library. For
annotation of identified clones (BLASTX, BLASTN and mapping to E.
coli K12 genome) please see French-Constant et al. 2000, Nucleic
Acids Res.
Seq primer: M13 Forward
Class: shotgun.

FEATURES	source
Location/Qualifiers	1..724
/organism="Photornabodus luminescens"	
/mol_type="genomic DNA"	
/strain="W14"	
/db_xref="taxon:29488"	
/clone="PLG01794"	
/dev_stage="primary phase variant"	
/clone_lib="Photornabodus luminescens strain W14 M13 library"	
/note="Genomic DNA from strain W14 was size selected (1.2 kb) and then cloned into M13 Janus."	
	218 a 176 c 137 g 192 t 1 others

	Query Match Best Local Matches	Similarity 242;	Score 12.0%;	Pred. 60.8%;	No.5-29; 0;	Mismatches 150;	Indels 5;	Gaps 1;
QY	118	ATGCTAAACCCCAAGTACTATCCTTAGCCACAGGAGGCACCATCGTGTTGGGGGA	177					
Db								
QY	404	ATTGCTTTACCAAATATTAACAATTITGGGGAGGGGGGACATATTCGGGTGCCGGAAG	345					
Db								
QY	178	TCTACGTCACAGATGAGCTACTCTCTGTGGAGCATCACCGTTGATPAAGCTTCTTGACCC	237					
Db								
QY	344	TCTGAAATCCCATCCAAATTAATGTGCTGGAAGGCTTGATATGATGAATTTAGTCAAATCT	285					
Db								
QY	238	GTCCCTGCCATCAACGACCTAGCCACCATCAAGGGGTGAAGAATCTCAAGATGGGCTCC	297					
Db								
QY	284	GTTCCTGAGATCAAAAATAATGCGCCAACCTAAAGGTGAGCAGATGATTAATTTGGTTGG	225					
Db								
QY	298	CAAAGAGGAGGGGGAAGSTGTGCTTAATATAGCCAAGCCTGTCAATAGACTCTCTGCC	357					
Db								
QY	224	CAGATATGAATGATACAGGTGTGGTTGACGCTTACCGAAMAAAMATCAACGCTGATTTGGAC	165					
Db								
QY	358	CAAAAAGAGACCGAGCGGTGATGATACCCATGSACTGACACATGCAAGGAAGAACCGCT	417					
Db								
QY	164	AAA-----ACCGATGGTTTTGTTATATACCACGSGTACTGTGACACAGTGAAGAAACCGCT	111					
Db								
QY	418	TTCTTCTCAACCTCACAGGTGAANAAGCCAAAAACCTGTGCCCTTGTAGAGCGCATCGT	477					
Db								
QY	110	TACTTCTCGATATCACTACCACTGCGCAAGCCGATATGATCAATGGTTGGCGCATGCC	51					
Db								
QY	478	CAAGGCTTTCATAGATGATGATGGGCCCCATGATATCT	515					
Db								
QY	50	CCATCTACGGCGCTGTGGTGGAGTGGGCCCGCTTAATCT	13					
Db								

RESULT 3	
A0500605/c	
LOCUS	A0500605
DEFINITION	V36C mtn-3kHA/lacZ Insertion Library Saccharomyces cerevisiae
ACCESSION	genomic 5', genomic survey sequence.
VERSION	A0500605
KEYWORDS	A0500605.1 GI:4706315
SOURCE	GSS.
ORGANISM	Saccharomyces cerevisiae (baker's yeast)
	Saccharomyces cerevisiae
	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;

REFERENCE 1 (bases 1 to 766)
 Saccharomyces; Saccharomycetaceae; Saccharomyces.
 Ross-Macdonald, P., Roemer, T., Coelho, P. S. R., Agarwal, S., Kumar, A.,
 deSeliges, S. A., Cheung, K. H., Sheehan, A., Symoniatzis, D., Jansen, R.,
 Umansky, L., Heidtman, M., Nelson, K., Iwasaki, H., Kanada, D., Lugo, R.,
 Hager, K., Miller, P., Roeder, G. S., and Snyder, M.
 Large-Scale Analysis of the Yeast Genome by Transposon Tagging and
 Gene Disruption

JOURNAL
 COMMENT
 Unpublished
 Contact: Kumar A
 Michael Snyder, Dept. of Mol. Cell. and Dev. Biology
 Yale University
 P.O. Box 208103, New Haven, CT 06520-8103, USA
 Tel: 203 432 9949
 Fax: 203 432 6161
 Email: anuj.kumar@yale.edu
 te of mtn-3kHA/lacZ insertion.
 Seq primer: GGCCTTCCTTCCTTGGAGTAC
 Class: transposon-tagged.

FEATURES
 source
 1..766
 /organism="Saccharomyces cerevisiae"
 /mol_type="genomic DNA"
 /db_xref="taxon:4932"
 /lab_host="E. coli"
 /clone_lib="mtn-3kHA/lacZ Insertion Library"
 /note="Vector: pHS6-Sal; A yeast genomic DNA library
 (lacking mitochondrial DNA) was prepared in pHS6-Sal;
 genomic DNA was size-fractionated (DNA of roughly 2-3 kb
 in length) prior to cloning. This library was
 subsequently mutagenized with a mtn-3kHA/lacZ
 multitransposon containing lacZ, URA3, and tet resistance."

BASE COUNT 223 a 175 c 168 g 198 t 2 others

ORIGIN
 Query Match 10.3%; Score 116.8; DB 28; Length 766;
 Best Local Similarity 50.2%; Pred. No. 2.9e-23;
 Matches 289; Conservative 0; Mismatches 287; Indels 0; Gaps 0;

QY 122 CTAACCCCAAGTACTATCTTACGACAGGAGGACCATGCTGGGGAATCTA 181
 |||||
 DB 628 CTTTGCCATCAATCAAAATTTTGGCCGCGGTACTTCCCTTCCAAAGTTGCAACA 569
 |||||
 QY 182 GCGTCAAGAGTACTCTGCTGGAGAGAGTCAAGCTTGAAGCTTTCGACGCTCC 241
 |||||
 DB 568 GGGCTACAAAGCGGGGTTATAGCGGCAATTACCGTATATATTAAAGAACGCTCC 509
 |||||
 QY 242 CTGCCATCAAGACCTAGCCACCAATCAAGGGTGAACAGATCTCAAGCATGTGCTCCCAAG 301
 |||||
 DB 508 CATCTTTAGCTGAGAGGACAGATCTGACTATCTTCAAGTGTCTAAGCTTGGTTCACAT 449
 |||||
 QY 302 AGATAGCGGTAGTAAAGTGGCTTAACTAGTCAAGGCTGCAATGCTCGCCCAAA 361
 |||||
 DB 448 CTTTAACTATAGCATCTGATCCATCTTATACAGGTATCTCCGAGGACATAGCTCTG 389
 |||||
 QY 362 AAGAGACGAAAGCCGTATCATACCATGAGTACCAATGAGAGAGACCGCTTCT 421
 |||||
 DB 388 ATGATTAGCGTGGGGTGTACCTGACGACGACCTATGAGGAGACAGCTTCT 329
 |||||
 QY 422 TCCTCAACCTACGCGTGAAGACCAAAACCTGTCTGCTTGTAGGCGCATGCTCCAG 481
 |||||
 DB 328 TCTTGATTGACCAATAATTCAGAGAACCCAGTATGTCCAGGCGCTATGCGCCAG 269
 |||||
 QY 482 GCTTCATCATGATGATGAGCCCATGATCTCTAATAGCGCGATGATGAGCATCA 541
 |||||
 DB 268 CCATGCGCACGCTGCTGATGCGCCCATATATTTATATAGCAGTGTCTATGCTGCT 209
 |||||
 QY 542 ACAAGCCCTCTACTAACAAGAGTGTGATGTGATGAGCATGATGATCAAGCCGCA 601
 |||||
 DB 208 CTGAGAAATACATGGGTGCTGGAGATGATCACTTAACAGTCTGATGCTGGGT 149
 |||||
 QY 602 GAGAACGACCAAGCTCAACACCGACGCTCATCATTTGCTTGGCCCAACAGGTA 661
 |||||

DB 148 TTGGACAACGAAATATGCAATCTTTAGATACATTCAGAGCGATGAACAGGAT 89
 QY 662 AAATCGCACAGTCTATTATGCAAGTGCAGTATT 697
 |||||
 DB 88 ATTAGGTTACTTTTCAATATGATGACGTCGAGTTT 53
 |||||

RESULT 4
 BH376509/c
 LOCUS
 DEFINITION AG-ND-137P15.TF.1 ND-TAM Anopheles gambiae genomic clone
 AG-ND-137P15, genomic survey sequence.
 ACCESSION BH376509
 VERSION BH376509.1 GI:17322651
 KEYWORDS GSS.
 SOURCE Anopheles gambiae (African malaria mosquito)
 ORGANISM Anopheles gambiae
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae;
 Anophelinae.

REFERENCE 1 (bases 1 to 757)
 Hong, Y. S., Hogan, J. R., Wang, X., Sarkar, A., Sim, C., Loftus, B. J., Ren,
 C., Huff, E. R., Carlile, J. L., Black, K., Zhang, H.-B., Gardner, M. J.,
 and Collins, F. H.
 Construction of a BAC library and generation of BAC end
 sequence-tagged connectors for genome sequencing of the African
 malaria mosquito Anopheles gambiae
 Mol. Genet. Genomics 268 (6), 720-728 (2003)
 22542063
 MEDLINE 1265338
 PUBMED

JOURNAL
 COMMENT
 Other GSSs: AG-ND-137P15.TF.1
 Contact: Brendan J Loftus
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 0208
 Fax: 301 838 3543
 Email: b.loftus@tigr.org
 This clone is from an A. gambiae BAC library (ND-TAM) provided by
 F. H. Collins and sequenced by The Institute for Genomic Research
 (TIGR). The BAC library was generated from A. gambiae PEST strain
 DNA. All DNA was extracted from newly hatched first instar larvae
 to minimize the inclusion of DNA from microorganisms that inhabit
 the gut. The DNA is derived from mixed sexes of larvae. The BAC
 library was constructed at Texas A&M University BAC Center
 University, College Station, Texas 77843-2123, USA using a HindIII
 partial digest.
 Seq primer: M13 For
 Class: BAC ends.

FEATURES
 source
 1..757
 Location/Qualifiers
 /organism="Anopheles gambiae"
 /mol_type="genomic DNA"
 /strain="PEST"
 /db_xref="taxon:7165"
 /clone_lib="AG-ND-137P15"
 /clone_lib="ND-TAM"
 /note="Vector: pECBAC1; Site_1: HindIII"

BASE COUNT 182 a 196 c 192 g 187 t

ORIGIN
 Query Match 9.8%; Score 110.8; DB 28; Length 757;
 Best Local Similarity 50.4%; Pred. No. 1.8e-21;
 Matches 271; Conservative 0; Mismatches 267; Indels 0; Gaps 0;

QY 204 TGGAGAGTCACCGCTGATATAGCTTCTGTAGCGCCGCCGACATCAAGACCTAGCCAC 263
 |||||
 DB 643 TGGAGCATCTTGGCCGATATCTTGGCTAGCGGCTGATGTAGGCAAGTACGCCAA 564
 |||||
 QY 264 CATCAAGGGGAGACAGATCTCAAGCATTTGGCTCCCAAGAGATGACGGGTAAAGTGGCT 323
 |||||
 DB 583 GATCCAGGTAAACACTTATTCACAGCTTCCATGAGCTAGTGGGCTGAGGCGTGGGT 524
 |||||

OY 324 TAAACCTAGCCAGCGTGTCAATGAGCTCTCCGCCAAAAAGAGACCGGATCAT 383
 Db 523 CCAGCTCAGCTGCTGTCTGCAAGAGCGTATCAACAGTCAAGCTGCTGGCGTATCGT 464
 OY 384 CACCCATGGAACCTGACACCAATGAGAGACCGCTTCTCTCCACACCTCCAGCTGAAAAG 443
 Db 463 TTCTCATGCGACAGATGCTGAGAGAAACAGCTTTTGGCTGTATGACAGTGAAGTC 404
 OY 444 CAAAAACCTGTCTGAGCGCCGATGCTCCAGGCTTTCATGAGTGTATG 503
 Db 403 CAATTAACCAATGCTGATTTGGCGACAGCAAGCCCTCGTTTCAAGCTTTGACGG 344
 OY 504 CCCCATGATCTCTATTAACCGCGGATGTAGCATCAACAAGCTTACTAACAAG 563
 Db 343 TCCGCGCACTGCTGATGCTGCGCATGCTTATCCCATCAAGATCGTGG 284
 OY 564 AGTGTGATTTGATGAGATGATTCAGCGCCGAGAGAGGACCAAGCTCAAC 623
 Db 283 CGCATGCTGCGATGAAACCAACGATCAACGCTACAGATGTACACAGACACAC 224
 OY 624 CACCGCATGATGATTTGCTTCCGCCACACAGGTAATAATGCGACAGTCTATTATG 683
 Db 223 TCCCAACGTTAAACTTTAAGTCTGATGATGCTGGCTTTATGATGAGTCTACCCAGA 164
 OY 684 CAAGTGTGATTTTCACTCAATCCGTTCCGACCTCACACCTTGAAGTGTATG 741
 Db 163 CAAGTGTGATTTTCACTCAATCCGTTCCGACCTCACACCTTGAAGTGTATG 106
 RESULT 5 843 bp DNA linear GSS 29-APR-1999
 AOS00657
 LOCUS 534F11 mtm-3xHA/lacZ Insertion Library Saccharomyces cerevisiae
 DEFINITION genomic 5', genomic survey sequence.
 ACCESSION AOS00657
 VERSION AOS00657.1 GI:4706367
 KEYWORDS GSS.
 SOURCE Saccharomyces cerevisiae (baker's yeast)
 ORGANISM Saccharomyces cerevisiae
 REFERENCE Saccharomyces cerevisiae
 AUTHORS Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 1 (bases 1 to 843)
 Ross-Macdonald, P., Roemer, T., Coelho, P. S. R., Agarwal, S., Kumar, A.,
 Desferes, S. A., Cheung, K. H., Sheehan, A., Symoniatidis, D., Jansen, R.,
 Umansky, L., Heideman, M., Nelson, K., Iwasaki, H., Kanada, D., Lugo, R.,
 Hager, K., Miller, P., Roeder, G. S., and Snyder, M.
 Large-Scale Analysis of the Yeast Genome by Transposon Tagging and
 Gene Disruption
 JOURNAL Unpublished
 COMMENT Contact: Kumar A
 Michael Snyder, Dept. of Mol. Cell. and Dev. Biology
 Yale University
 P.O. Box 208103, New Haven, CT 06520-8103, USA
 Tel: 203 432 9949
 Fax: 203 432 6161
 Email: anuj.kumar@yale.edu
 te of mtm-3xHA/lacZ insertion.
 Seq primer: GGCCTCTCTTCTTGGAGATAC
 Class: transposon-tagged.
 Location/Qualifiers
 1. 843
 /organism="Saccharomyces cerevisiae"
 /mol_type="genomic DNA"
 /db_xref="taxon:4932"
 /lab_host="E. coli"
 /clone_lib="mtm-3xHA/lacZ Insertion Library"
 /note="Vector: PHSS6-Sal; A yeast genomic DNA library
 (lacking mitochondrial DNA) was prepared in PHSS6-Sal;
 genomic DNA was size-fractionated (DNA of roughly 2-3 kb
 in length) prior to cloning. This library was
 subsequently mutagenized with a mtm-3xHA/lacZ
 multitransposon containing lacZ, URA3, and tet resistance."
 BASE COUNT 245 a 182 c 182 g 224 t 10 others

ORIGIN
 Query Match 9.6%; Score 108.6; DB 28; Length 843;
 Best Local Similarity 50.7%; Pred. No. 8.6e-21;
 Matches 255; Conservative 0; Mismatches 248; Indels 0; Gaps 0;
 OY 208 GCAGTCACCGTGTATGATCTTCTTGGACCGCTCCCTCCATCAACGCTAGCCACCATC 267
 Db 679 GATTAACCGTAAATGATATGATGATGAGAGCGGTGCATCTTATAGTGAGAGGACATCTG 620
 OY 268 AAGGAGACAGATCTCAACATTTGGCTCCCAAGAGATGACGGGTAAAGTGTGGCTTAAA 327
 Db 619 GACTATCTTCAATGCTTAACGTTGCTTCAATTTCTTAAACATAGCATATCATCCCA 560
 OY 328 CTAGCCAGAGGTGTCAATGAGCTCTCCGCCAAAAAGAGACGAGCGGATGATCACC 387
 Db 559 TTGTATCAGCGTATCTTCGAGGACGATGCTTGTATGATACATGCTGTGGTGTGCTACT 500
 OY 388 CATGAACTGACACCATGAGAGACCGCTTCTCTCAACCTCAGCGGTGAAAAAGCCAA 447
 Db 499 CATGAGACGACACTATGAGAGACAGCTTCTCTTATGATTTGACCATTAATTTCAAG 440
 OY 448 AAACCTGTGCTTGTATGAGCGGATGCGCTCCAGGCTTCCATGATGATGCGCCC 507
 Db 439 AAGCCAGTATGATGAGCGGCTATGCGTCCAGCCACTCCAGCTGCTGATGAGCCCA 380
 OY 508 ATGAATCTCTATTAACCGCGGATGATGATGATCAACAAACCTCTACTAACAAGAGATG 567
 Db 379 ATGAATTTATATCAACAGAGTGTATTTGCTTGTGAGAAATACATGCGTGGCGACG 320
 OY 568 GTGATTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 627
 Db 319 ATGATCACTCTAATCAACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 260
 OY 628 GCAGTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 687
 Db 259 TCTTTGATACATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 200
 OY 688 GTGAGTATTTCACTCAATCCGT 710
 Db 199 GTGAGTATTTCACTCAATCCAGT 177
 RESULT 6 768 bp DNA linear GSS 17-DEC-2002
 BZ570599/c
 LOCUS msh2_1446.y2 msh Pseudomonas aeruginosa genomic clone msh2_1446,
 DEFINITION genomic survey sequence.
 ACCESSION BZ570599.1 GI:27205660
 KEYWORDS GSS.
 SOURCE Pseudomonas aeruginosa
 ORGANISM Pseudomonas aeruginosa
 REFERENCE Pseudomonadaceae: Pseudomonas.
 1 (bases 1 to 768)
 Spencer, D. H., Raymond, C. K., Smith, E. E., Sims, E. E., Hastings, M.,
 Burns, J. L., Kaul, R., and Olsen, M. V.
 Whole-Genome-Sequence Variation among Multiple Isolates of
 Pseudomonas aeruginosa Library
 J. Bacteriol., (2002) In press
 Contact: Chris. K. Raymond
 Genome Center
 University of Washington
 Box 352145, Seattle, WA 98105-2145, USA
 Tel: 2062216954
 Fax: 2066857244
 Email: craymond@u.washington.edu
 Class: Shotgun.
 Location/Qualifiers
 1. 768
 /organism="Pseudomonas aeruginosa"
 /mol_type="genomic DNA"

[illegible]

Db	109	TAACAGATGTATVGGCTCTGGGTTTGGACAAACGAATGAATGCCAACCTTTAGATA	50
QY	638	CATT 641	
Db			
	49	CATT 46	
RESULT 13			
LOCUS	AO503362/c		
DEFINITION	AO503362	561 bp	linear
ACCESSION	V57D10	mtn-3xHA/lacZ	Insertion Library
VERSION	AO503362	genomic 5',	genomic survey sequence.
KEYWORDS	AO503362.1	GI:4705908	
SOURCE	SS.		
ORGANISM	Saccharomyces cerevisiae	(baker's yeast)	
REFERENCE	Saccharomyces cerevisiae		
AUTHORS	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes; Saccharomycetales; Saccharomycetaceae; Saccharomyces.		
	1 (bases 1 to 561)		
	Ross-Macdonald, P., Roemer, T., Coelho, P.S.R., Agarwal, S., Kumar, A.,		
	destages, S.A., Cheung, K.-H., Sheehan, A., Symonidis, D., Jansen, R.,		
	Umansky, L., Heidman, M., Nelson, K., Iwasaki, H., Kanada, D., Lugo, R.,		
	Hager, K., Miller, P., Roeder, G.S. and Snyder, M.		
	Large-Scale Analysis of the Yeast Genome by Transposon Tagging and		
	Gene Disruption		
TITLE	Unpublished		
JOURNAL	Contact: Kumar A		
COMMENT	Michael Snyder, Dept. of Mol. Cell. and Dev. Biology		
	Yale University		
	P.O. Box 208103, New Haven, CT 06520-8103, USA		
	Tel: 203 432 9949		
	Fax: 203 432 6161		
	Email: anuj.kumar@yale.edu		
	te of mtn-3xHA/lacZ insertion.		
	Seg primer: GGCCTTCTCTTGGAGACTAC		
	Class: transposon-tagged.		
FEATURES	Location/Qualifiers		
source	1..561		
	/organism="Saccharomyces cerevisiae"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:4932"		
	/lab_host="E. coli"		
	/clone_lib="mtn-3xHA/lacZ Insertion Library"		
	/note="Vector: pHS56-Sal; A yeast genomic DNA library		
	(lacking mitochondrial DNA) was prepared in pHS56-Sal;		
	genomic DNA was size-fractionated (DNA of roughly 2-3 kb		
	in length) prior to cloning. This library was		
	subsequently mutagenized with a mtn-3xHA/lacZ		
	minitransposon containing lacZ, URA3, and tet resistance."		
BASE COUNT	158 a 128 c 118 g 157 t		
ORIGIN			
Query Match	8.2%;	Score 93;	DB 28;
Best Local Similarity	53.4%;	Pred. No. 3e-16;	Length 561;
Matches	195;	Conservative 0;	Mismatches 170;
		Indels 0;	Gaps 0;
QY	346	GAGCTCTCGCCCAAAAAGAGACGAACCGGTGATCATCAACCATGAGACTGACCCGATG	405
Db	514	GAGCATTTAGCATTTGATGATTAGCGTGGCGGTTGTCATCTCATGGACGACGACATATG	455
QY	406	GAAAGACGCGTTTCTTCCTCCCAACCTCAGCTGTAAGAAACCAAAACCTGCTGCTTGTGA	465
Db	454	GAGAGACAGCTTTCTTCTTCTTATGATTTCACATAGATTTCAGAGAAAGCCAGTATGATTCGA	395
QY	466	GCGCCCATGGCTCCAGGCTCTTCATGATGATGATGATGATGATGATGATGATGATGATGATG	525
Db	394	GGCGCTATGGGTCGACGACATCGCAGCTGCTGATGATGATGATGATGATGATGATGATGATG	335
QY	526	GTGATGTAGCATCAACAAGCCTCTACTACAAAGAGAGTGTGATTTGATGAACGAT	585
Db	334	GTGCTATTGATGCTTCTGAGAAATCATCGTGGTTCGTCAGACGATGATCACTTAAACGAT	275

QY	586	GAGATTACGCGCCGACAGAGAGGCGACGACCAAGCCACACACCACCGAGTCATATGCATTCTCT	645
Db	274	CGTATTGCGCTCTGGGTTTGGACAACGAAATGATGCGCACTCTTTGATACATCTAGA	215
QY	646	TCGCCCAACACAGGTAAATCGCGACAGCTATATTATGGCAAAAGTCAGATATTACTCAA	705
Db	214	GGGATATGAACAGGGATATTATTAGTTACTTTCAATGATGAGCGACTTTTACTACCCA	155
QY	706	TCGGT 710	
Db	154	CCAGT 150	
RESULT 14			
BH371112/c			
LOCUS			
DEFINITION	AG-ND-137P15.TF ND-TAM Anopheles gambiae genomic clone AG-ND-137P15		
ACCESSION	BH371112	663 bp	DNA
VERSION	BH371112		linear
KEYWORDS	BH371112.1 GI:17317237		GSS 10-DEC-2001
SOURCE	GSS.		
ORGANISM	Anopheles gambiae (African malaria mosquito)		
REFERENCE	Anopheles gambiae		
AUTHORS	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Anopheles.		
TITLE	1 (bases 1 to 663)		
JOURNAL	Hong,Y.S., Hogan,J.R., Wang,X., Sarkar,A., Sim,C., Loftus,B.J., Ren		
MEDLINE	,C., Huff,E.R., Carille,J.L., Black,K., Zhang,H.-B., Gardner,M.J.		
PUBMED	and Collins,F.H.		
COMMENT	Construction of a BAC library and generation of BAC end sequence-tagged connectors for genome sequencing of the African malaria mosquito Anopheles gambiae		
	Mol. Genet. Genomics 268 (6), 720-728 (2003)		
	12542063		
	12655398		
	Other-GSS: AG-ND-137P15.TR		
	Contact: Brendan J Loftus		
	Department of Eukaryotic Genomics		
	The Institute for Genomic Research		
	9712 Medical Center Dr., Rockville, MD 20850, USA		
	Tel: 301 838 0208		
	Fax: 301 838 3543		
	Email: b1o@fusedti.gr.org		
	This clone is from an A. gambiae BAC library (ND-TAM) provided by F.H. Collins and sequenced by The Institute for Genomic Research (TIGR). The BAC library was generated from A. gambiae PEST strain DNA. All DNA was inactivated from newly hatched first instar larvae to minimize the inclusion of DNA from microorganisms that inhabit the gut. The DNA is derived from mixed sexes of larvae. The BAC library was constructed at Texas A&M University BAC Center University, College Station, Texas 77843-2123, USA using a HindIII partial digest.		
	Seq primer: M13 For		
	Class: BAC ends.		
FEATURES			
SOURCE	Location/Qualifiers		
	1..663		
	/organism="Anopheles gambiae"		
	/mol_type="genomic DNA"		
	/strain="PEST"		
	/db_xref="taxon:7165"		
	/clone="AG-ND-137P15"		
	/clone_lib="ND-TAM"		
	/note="Vector: pECBAC1, site_1: HindIII"		
BASE COUNT	160 a 175 c 168 g 160 t		
ORIGIN			

	Query Match	7.9%	Score 89.6;	DB 28;	Length 663;
	Best Local Similarity	50.4%;	Pred. No. 3.4e-15;		
	Matches 244;	Conservative 0;	Mismatches 13;	Indels 1;	Gaps 1;
QY	259	GCACCAATCAAGGCTGAACACATCTCAACGACTTGGCTCCCAAGATGACGGTAAAGGTG	318		

Db	589	GCCAGATCCAGGATTAACAACTTATCCAAACGTTCCATCGGATACATATGGGCGCTACGGC	530
QY	319	TGGCTTAAACTGAGCCAGCGGTGTCAATGACTCTCGGCCAAAAAGACGGAACCGGTG	378
Db	529	TGGGTCCAGCTACATGCTGTGTGGCAAGCCGTATCAACAGTCGACGTTGCTGGCGTG	470
QY	379	ATCATCACCACATGGAAGTGCACCACTGGAAGAGACCGGCTTTGTTCTCAACCTCACGGTG	438
Db	469	ATCGTTTCTCATGGCACAGATACGCTGGAGGAAAGCTTTTGGCTTGATTTGACAGTG	410
QY	439	AAAAGCCAAAACCTGTCTGCCCTTTAGCGGCCATGCGTCCAGGCTCTTCCATGAGTCT	498
Db	409	AAGTCCAATTAACCAATGCTGTCTGATTTGGCGCACAGGCCCAAGCGCTCGGTTTCAACACTTT	350
QY	499	GATGCCCCCATGAACTCTCATTAAGCCGCGTGAATGAGGATCAACAAAGCCCTACTAC	558
Db	349	GACGGTCCGGCGCACTTGCTTAAGCCTGTGCGCATGCGTGTGATGCCCAATCAAGGAT	290
QY	559	AAAAGAGTGATGTTGTATGAAGCATGATTTACGGCCGCGAGGAAGCGACCAAGCTC	618
Db	289	CGTGGCGGCATGCTGGCATGCATGAACCAACGATACAGCTCAAGTATGTGCACCAAGCA	230
QY	619	AACACACCCGAGTAATGCAATT--GCTTCGCCCAACACAGGTAAATACGGCAGCTCA	677
Db	229	CACACTGCCAAGTTGAACCTTTTAAGTTGTGGTGAATGGCGCTTTATTGGTAGGCTA	170
QY	678	TTATGGCAAGTGCAGATATTCATCAATCCGTTTGAGCTACACACCTTGAAGAGCTT	737
Db	169	CCAGACACAAGTCGTTTACCGACTAGGCGCATCCGCAATTTAACTTCTCATTAAGC	110
QY	738	TGAT 741	
Db	109	AGAT 106	

FEATURES	source
LOCUS	AO873892/c
DEFINITION	AO873892 560 bp DNA linear GSS 08-NOV-1999
ACCESSION	V8345 mtN-3xHA/lacZ insertion library, strain AB972 Saccharomyces cerevisiae genomic 5', genomic survey sequence.
VERSION	AO873892
KEYWORDS	AO873892.1 GI:6286136
SOURCE	GSS.
ORGANISM	Saccharomyces cerevisiae (baker's yeast)
REFERENCE	Saccharomyces cerevisiae
AUTHORS	Eukaryota, Fungi; Ascomycota; Saccharomycotina; Saccharomycetes; Saccharomycetales; Saccharomycetaceae; Saccharomyces.
TITLE	1 (bases 1 to 560)
JOURNAL	Ross-Macdonald,P., Roemer,T., Coelho,P.S.R., Agarwal,S., Kumar,A., deEsteygas,S.A., Cheng,K.-H., Sheehan,A., Symonitis,D., Jensen,R., Umansky,L., Heldman,M., Nelson,K., Iwasaki,H., Kanada,D., Lugo,R., Hager,K., Miller,P., Roeder,G.S. and Snyder,M.
COMMENT	Large-Scale Analysis of the yeast Genome by transposon Tagging and Gene Disruption
UNPUBLISHED	Contact: Kumar A
UNPUBLISHED	Michael Snyder, Dept. of Mol. Cell. and Dev. Biology
UNPUBLISHED	Yale University
UNPUBLISHED	P.O. Box 208103, New Haven, CT 06520-8103, USA
UNPUBLISHED	Tel: 203 432 9949
UNPUBLISHED	Fax: 203 432 6161
UNPUBLISHED	Email: annj.kumar@yale.edu
UNPUBLISHED	te of mtN-3xHA/lacZ insertion.
UNPUBLISHED	Seq primer: GGCTCTCTCTTGTGGAAGTAC
UNPUBLISHED	Class: transposon-tagged.
UNPUBLISHED	Location/Qualifiers
UNPUBLISHED	1..560

```

/organism="Saccharomyces cerevisiae"
/mol_type="genomic DNA"
/strain="AB972 - tripl r(0) (S288C background)"
/db_xref="taxon:4932"
/lab_host="E. coli"
/clone_lib="trn-3xHA/lacZ insertion library, strain AB972

```

/note="Vector: pHS6-Sal: A yeast genomic DNA library was prepared in pHS6-Sal; genomic DNA was size-fractionated (DNA of roughly 2-3 kb in length) prior to cloning. This library was subsequently mutagenized with a min-3xHA/lacZ minitransposon containing lacZ, URA3, and tet resistance.

BASE COUNT 155 a 127 c 119 g 158 t 1 others
ORIGIN

Query Match 7.7%; Score 87; DB 28; Length 560;

Best Local Similarity 54.0%; Pred. No. 1.8e-14;

Matches 177; Conservative 0; Mismatches 151; Indels 0; Gaps 0;

```
QY 383 TCACCCATGGAGACGACACATGGAGAGACCGCTTCTCCTCAACCTCAGCGTGAAA 442
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 503 TCACTCATGGAGACGACACTATGGAGAGACAGCTATTCTTAGATTGACCATANATT 444
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 443 GCCAAAAACCTGTCTGCCCTGTAGGCCCATGCTCCAGGCTCTTCATGAGTGTGATG 502
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 443 CAGAGAAGCCAGTATGATCGACGCGATATGCTCCAGCCACTGCCAGCTGCTGTGATG 384
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 503 GCCCATGATCTCTATACGCCCTGATATGATGATGATGATGATGATGATGATGATG 562
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 383 GCCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 324
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 563 GAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 622
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 323 GCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 264
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 623 CCACCGCAGTCAATGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 682
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 263 CCAACTCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 204
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 683 GCAAGTCGAGTATTTCACTCAATCCGT 710
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 203 ATGACGTGAGTTTACTACCACCACT 176
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

Search completed: September 4, 2003, 18:13:36
Job time : 2342 secs